UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE BIOLOGIA VEGETAL

A mouse model of infection to investigate Kaposi's sarcoma-associated herpesvirus LANA function *in vivo*

Nelly Marine Carreira da Silva

Mestrado em Biologia Molecular e Genética

Dissertação orientada por:
Professor Doutor João Pedro Monteiro e Louro Machado de Simas

Professora Doutora Maria Filomena Ribeiro Alcobia da Silva Trabucho Caeiro

2018

Acknowledgements

I would like to thank Professor Dr. Pedro Simas, my supervisor at Instituto de Medicina Molecular, for accepting me in his lab, for all the good advices and for believing in me and in my competences. I want to thank to Professor Dr. Filomena Caeiro, my internal supervisor, for her motivation, advices and questions, and for her availability. Thanks to Faculdade de Ciências (UL) for giving me the opportunity to attend Biologia Molecular e Genética master's degree.

I would also like to thank to all the members that belonged or belong to PSimas lab, specially to Marta for all the support, help and advices she gave me during this journey.

To all my friends, thank you for being supportive. A special thanks to Daniela and Mariana for their kind words.

Lastly, I want to thank to my family. They have always done their best to support and help me.

Abstract

Herpesviruses are ubiquitous in the world infecting a wide range of hosts. Eight herpesviruses have been related to humans, including Kaposi's sarcoma-associated herpesvirus (KSHV). KSHV is the etiological agent of several malignancies, namely, Kaposi's sarcoma (KS), primary effusion lymphoma (PEL) and multicentric Castleman's disease (MCD). As other members of *Herpesviridae* family, KSHV displays a biphasic life cycle constituted by a lytic and a latent phase. The latent phase has a highly restricted gene expression and the viral genome is circularized into an episome which is maintained in the host's B cells. Its narrow host range constitutes a major limitation for *in vivo* studies. Murine herpesvirus 68 (MHV-68), which is genetically related to KSHV, is able to infect laboratory mice and to establish a latent persistent infection. The principal protein related to this phase is the latency-associated nuclear antigen (LANA). LANA from KSHV (kLANA) and LANA from MHV-68 (mLANA) are functionally related and both have been shown to be required for the maintenance of latency in KSHV and MHV-68, respectively. Recently, a MHV-68 chimeric virus (v.kLANA) was constructed where *mLANA* was replaced by *kLANA*. This constitutes a suitable model to study kLANA functions *in vivo*.

In vitro studies demonstrated that the region encompassing amino acids 33 to 194 of kLANA is crucial for segregation. This function, together with replication, is critical for the maintenance of the viral episome. Therefore, the aim of this study was to assess the importance of that region using an *in vivo* model of infection. This was achieved by generating a chimeric recombinant virus where *mLANA* was replaced by *kLANA* containing a deletion encompassing amino acids 33 to 194 within MHV-68 genome.

Our *in vitro* assays showed that recombinant viruses were able to express the LANA mutant protein and had similar growth kinetics to MHV-68 (v.WT) and v.kLANA. The results from *in vivo* experiments demonstrated that v.kLANAΔ33-194.yfp had similar levels of latency and frequency of viral DNA positive cells to v.kLANA.yfp. However, v.kLANAΔ33-194 showed significant lower latency levels compared to v.kLANAΔ33-194.yfp and v.kLANA.yfp. Since the two recombinant viruses had discordant phenotypes, the importance of the region remains unclear. Further studies need to be done to clarify the importance of the region.

Key-words: KSHV; MHV-68; LANA; chimeric virus.

Resumo

Os herpesvírus são vírus ubíquos que infetam uma ampla variedade de hospedeiros. Até ao momento, oito herpesvírus foram relacionados com o ser humano, nomeadamente, o herpesvírus associado ao sarcoma de Kaposi (KSHV). O KSHV é o agente etiológico de várias doenças, inclusive do Sarcoma de Kaposi e de outras doenças linfoproliferativas tais como o linfoma de efusão primária (PEL) e a doença de Castleman multicêntrica (MCD). Comummente com os outros herpesvírus, o KSHV possui um ciclo de vida bifásico constituído por uma fase lítica e uma fase latente. A fase lítica é caracterizada pela expressão dos genes virais e produção de novos viriões que, consequentemente, levam à lise celular. Na fase latente existe uma repressão da expressão génica e o genoma viral é circularizado, mantendo-se sob a forma de episoma no núcleo das células B do hospedeiro. O balanço entre as fases lítica e latente é essencial para a manutenção e sobrevivência do vírus. O genoma do KSHV tem um tamanho aproximado de 140 kpb e é constituído por uma região longa única, que inclui as regiões codificantes e se encontra flanqueada por um número variável de repetições terminais (TR). Uma das grandes limitações dos estudos *in vivo* deste vírus deve-se à sua especificidade de hospedeiro, uma vez que só infecta humanos. O herpesvírus murino 68 (MHV-68), com um genoma de 118 kpb, está geneticamente relacionado com o KSHV e consegue estabelecer infeção latente em murganhos (*Mus musculus*). Após a inoculação intranasal do vírus em murganhos, ocorre uma fase de replicação lítica nas células epiteliais pulmonares. De seguida, o vírus dissemina-se para o baço onde infecta as células B do centro germinativo. Nesse local ocorre a proliferação de células B infetadas, sucedendo-se o estabelecimento de latência nas células B de memória. Estas constituem o seu reservatório a longo prazo.

Apesar de haver uma elevada restrição da expressão génica durante a fase latente do ciclo de vida do KSHV, alguns genes são expressos, nomeadamente os que codificam o antigénio nuclear associado à latência (LANA), a v-ciclina, a proteína v-FLIP e as *kaposins*. A LANA do KSHV (kLANA) é a principal proteína expressa ao longo desta fase e a sua região N-terminal interage com as histonas H2A e H2B, o que permite a sua associação aos cromossomas do hospedeiro. A região C-terminal contém um domínio de ligação ao DNA (DBD) que se liga a regiões específicas nas TR do genoma viral. Estas interações são ambas importantes para a replicação e segregação do episoma, funções cruciais para a persistência viral. A proteína LANA do MHV-68 (mLANA), semelhantemente à kLANA, contém um DBD que interage com as TR. Além disso, também interage com os cromossomas mitóticos. A interação entre o genoma viral do MHV-68 e o genoma do hospedeiro processa-se possivelmente através de um mecanismo semelhante ao mecanismo adotado pelo KSHV. Portanto, kLANA e mLANA são proteínas estruturalmente e funcionalmente conservadas e possivelmente relacionadas. Ambas mostraram ser essenciais para o estabelecimento e manutenção da latência do KSHV e MHV-68, respetivamente.

Recentemente, foi desenvolvido um vírus quimera (v.kLANA) em que a *kLANA* foi clonada no genoma do MHV-68, substituindo a *mLANA* endógena. Este vírus quimera consegue estabelecer a fase latente em murganhos, apesar de se observarem níveis de latência mais baixos comparativamente aos níveis obtidos com o MHV-68. Deste modo, o vírus quimera constitui um bom modelo para estudar as funções de kLANA *in vivo*.

Estudos *in vitro* recentes concluíram que a região entre os aminoácidos 33 e 194 da kLANA é essencial para a segregação do episoma. Esta função e a replicação são críticas para a manutenção do episoma do KSHV. Considerando estes resultados, o objetivo deste estudo consistiu em avaliar a importância dessa região no estabelecimento de latência *in vivo*, utilizando o modelo quimera supramencionado. Para esse fim, construiu-se um vírus quimera recombinante com o gene que codifica

a proteína kLANA contendo uma deleção entre os aminoácidos 33 a 194 (*kLANAΔ33-194*). Esse gene foi clonado num cromossoma artificial bacteriano (BAC) que contém o genoma do MHV-68, substituindo o gene que codifica a mLANA. Os vírus recombinantes foram construídos em dois *backgrounds*, um em que o vírus expressa a *yellow fluorescent protein* (YFP) e outro em que não expressa esta proteína. A expressão da proteína YFP é vantajosa pela possibilidade de rastrear a infeção através de análises por citometria de fluxo, por exemplo. Após a obtenção de vários clones, estes foram analisados através de reações de digestão com as enzimas de restrição *Eco*RI, *Bam*HI e *Hind*III. Dos que apresentaram as bandas esperadas nos perfis de digestão, foram selecionados dois clones, um de cada *background*, para prosseguir o estudo.

Os estudos *in vitro* realizados incluíram a análise da cinética de crescimento dos vírus e a expressão de proteínas. Para analisar a cinética de crescimento, células permissivas foram infetadas com os vírus recombinantes (v.kLANAΔ33-194 e v.kLANAΔ33-194.yfp) e com os vírus controlo (v.kLANA e MHV.68) de ambos *backgrounds*. As células infetadas foram recolhidas em seis tempos específicos da infeção (0, 24, 48, 72, 96 e 120 horas pós infeção). Após um processo de congelamento e descongelamento, as amostras foram tituladas. Os títulos obtidos para cada tempo permitiram traçar uma curva de crescimento para cada vírus. Verificou-se que os vírus recombinantes apresentaram uma cinética de crescimento semelhante aos controlos utilizados. Além desta experiência, também foi realizado um *western blot* para a deteção da proteína kLANA usando como controlos as proteínas virais mLANA, M3 (quimiocina) e *green fluorescent protein* (GFP) e a proteína celular actina. Este ensaio permitiu verificar que os vírus recombinantes expressaram a proteína kLANA mutante com uma massa molecular expectavelmente inferior à apresentada pela proteína kLANA *wild-type*. Para analisar a relevância da região no estabelecimento de latência *in vivo*, 4 grupos de 5 murganhos da estirpe C57BL/6J foram infetados com os vírus recombinantes (v.kLANAΔ33-194 e v.kLANAΔ33-194.yfp), usando como controlos MHV.68 (v.WT.yfp) e v.kLANA.yfp. Ao 14º dia após a inoculação intranasal (pico de latência), os murganhos foram sacrificados e os baços foram cirurgicamente removidos e devidamente processados para os ensaios a realizar. Para avaliar o nível de latência foi realizado um ensaio de reativação *ex vivo* e também se determinou a frequência de células positivas para a presença de DNA viral através do ensaio de diluição limitante e PCR em tempo real. Em ambas as experiências, v.kLANA.yfp e v. kLANAΔ33-194.yfp apresentaram níveis semelhantes, quer de latência, quer de células positivas para DNA viral. No entanto, verificou-se que nessas mesmas experiências, v.kLANAΔ33-194 apresentou níveis significativamente mais baixos de latência e de células positivas para a presença de DNA viral comparativamente com v.kLANAΔ33-194.yfp. A obtenção de um fenótipo discordante entre v.kLANAΔ33-194 e v.kLANAΔ33-194.yfp não era esperada dado que o acoplamento da proteína YFP ao genoma do MHV-68 mostrou não interferir com os níveis de latência. Foi também realizada uma análise por citometria de fluxo que permitiu quantificar a percentagem de células B do centro germinativo que estão infetadas, através da deteção da proteína YFP expressa pelos vírus v.WT.yfp, v.kLANA.yfp e v. kLANAΔ33-194.yfp . Na experiência realizada verificou-se que v.kLANAΔ33-194.yfp apresenta uma frequência de infeção de células B do centro germinativo ligeiramente inferior (aproximadamente 1%), comparando com v.kLANA.yfp. No entanto, a experiência terá de ser repetida de modo a obter resultados estatisticamente significativos.

Tendo em conta a discordância inesperada entre os fenótipos dos vírus recombinantes, a avaliação da importância da região em estudo utilizando o modelo vírus quimera não foi elucidada neste trabalho, sendo necessárias experiências futuras.

Palavras-chave: KSHV; MHV-68; LANA; vírus quimera.

Table of contents

1. Introduction **1**
1.1 Herpesviridae family 1
1.2. *Gammaherpesvirinae* subfamily 2
1.3. Kaposi's sarcoma associated herpesvirus (KSHV) 2
1.3.1. Lytic phase 4
1.3.2. Latent phase 4
1.3.3. kLANA 5
1.4. Murine herpesvirus 68 (MHV-68) 7
1.4.1. mLANA 8
1.5. Chimeric virus to study KSHV pathogenesis 9
2. Aim of the project 10
3. Materials and methods 10
3.1 Materials 10
3.1.1. Plasmids 10
3.1.2. Bacteria 11
3.1.3. Cell lines 11
3.1.4. Viruses 11
3.1.5. Mice 12
3.2. Methods 12
3.2.1 Construction of chimeric virus MHV-68-kLANAΔ33-194 12
3.2.1.1 Isolation and analysis of DNA 12
3.2.1.1.1. Restriction endonuclease reactions 12
3.2.1.1.2. Polymerase chain reaction (PCR) 12
3.2.1.1.3. Analysis and isolation of DNA by gel electrophoresis 12
3.2.1.1.4. DNA Ligation 12
3.2.1.1.5. DNA quantification and sequencing 12
3.2.1.1.6 DNA plasmid isolation – Miniprep 12
3.2.1.1.7. BAC DNA preps 13
3.2.1.2. Cloning procedures 13
3.2.1.2.1. Subcloning of kLANAΔ33-194 into pSP72_PCR1-5 13
3.2.1.2.2. Subcloning of pSP72_kLANAΔ33-194 insert into BamHI-G shuttle vector 13
3.2.1.3. Bacterial methods 14
3.2.1.3.1. Transformation of competent cells 14
3.2.1.3.2. Mutagenesis 14

3.2.2. Reconstitution of MHV-68_kLANAΔ33-194 14
3.2.2.1 Virus Reconstitution on BHK-21 14
3.2.2.2. Passage of virus through NIH3T3-Cre cells to remove BAC sequences 15
3.2.3.1. Production of viral stocks 15
3.2.3.2. Virus titration using suspension assay – plaque assay 15
3.2.3.3. Multi-step growth curve 15
3.2.3.4. Immunoblotting 15
3.2.4. Animal experiments 16
3.2.4.1 Ethics statement 16
3.2.4.2. Infection of mice 16
3.2.4.3. Single cell suspension 16
3.2.4.4. Infectious center assay (ICAs) 17
3.2.4.5. Limiting dilution assay and analysis by real-time PCR 17
3.2.4.6. Flow cytometry 17
3.2.4.7. Statistical analysis 18
4. Results 18
4.1. Generation and characterization of MHV-68 recombinant virus 18
4.2. *In vitro* assays 21
4.2.1. Viral titration 21
4.2.2. Expression of kLANA mutant proteins 21
4.2.3. Growth kinetics of recombinant viruses 22
4.3. *In vivo* assays 23
4.3.1. Infectious center assay 23
4.3.3. Flow cytometry analysis 26
5. Discussion and future perspectives 28
6. References 30
Supplementary Data 36

Figure index

Figure 1.1. Schematic diagram of a human herpesvirus. 2
Figure 1.2. KSHV episome 4
Figure 1.3. Representation of KSHV life cycle in an infected cell.. 5
Figure 1.4. Representation of full-length kLANA protein. 6
Figure 1.5. Mechanism of tethering used by kLANA 7
Figure 1.6. MHV-68 infection 8
Figure 1.7. Comparison of full-length kLANA and mLANA proteins……...9
Figure 1.8. Schematic diagram of the construction of v.kLANA. 9
Figure 4.1. Schematic representation of kLANA and mutant kLANA (kLANAΔ33-194) proteins…..18
Figure 4.2. Restriction profiles of BAC plasmids…….....19
Figure 4.3. Schematic representation of genome depicting each restriction enzyme's sites.….....20
Figure 4.4. Western Blot analysis of viral proteins (kLANA, mLANA, M3 and eGFP) and cellular protein (actin) in v.WT, v.kLANA and v.kLANAΔ33-19…….....22
Figure 4.5. Multi-step growth curve of v.WT and chimeric viruses, in both backgrounds…….....23
Figure 4.6. Quantification of latent infection (PFU/spleen) in the spleens of mice infected with v.WT.yfp, v.kLANA.yfp, v.kLANAΔ33-194.yfp and v.kLANAΔ33-94……..…24
Figure 4.7. Frequency of viral DNA positive cells in total splenocytes, obtained by limiting dilution assay and real-time PCR……25
Figure 4.8. Flow cytometry analysis. Percentage quantification of GC B cells (CD19+) and representative FACS dot plot gated on B cells……..26
Figure 4.9. Flow cytometry analysis. Percentage quantification of YFP positive (infected) GC B cells (CD19+GL7+CD95+) and representative FACS dot plot gated on GC B cells……27
Figure 4.10. Flow cytometry analysis. Percentage quantification of YFP positive B cells with GC phenotype (CD19+YFP+) and representative FACS dot plot gated on YFP positive B cells……...27

Table index

Table 1.1. Human hespesvirus and their key features.. 1
Table 4.1. Viral stocks titration.. 21
Table 4.2. Reciprocal frequency of viral DNA positive cells in total splenocytes..... 25

Abbreviations

BAC: Bacterial artificial chromosome

BHK: Baby hamster kidney

CPE: Cytopathic effect

CRISPR: Clustered regularly interspaced short palindromic repeats

CWS: Cell working stocks

DBD: DNA-binding domain

DBS: Double strand break

DC: Dendritic cell

DE: Aspartate- and glutamate-rich region

DMEM: Dulbecco's modified Eagle medium

DPI.: Days post-infection

E: Early

EBNA: Epstein-Barr nuclear antigen

EBV: Epstein-Barr virus ORF – open reading frame

EQE: glutamate- and glutamine-rich region

FACS: Fluorescence-activated cell sorting

FBS: Fetal bovine serum

FDS: Folicular dendritic cells

GC: Germinal center

GFP: Green fluorescent protein

HCMV: Human cytomegalovirus

HHV-6: Human herpesvirus 6

HHV-7: Human herpesvirus 7

HSV1: Herpes simplex virus type 1

HSV2: Herpes simplex virus type 2

HVS: Herpesvirus saimiri

ICA: Infectious center assay

IE: Immediate early

IRIS: KSHV immune reconstitution inflammatory syndrome

KICS: KSHV-associated inflammatory cytokine syndrome

KS: Kaposi's sarcoma

KSHV: Kaposi's sarcoma-associated herpesvirus

L: Late

LANA: Latency-associated nuclear antigen

LB: Luria-Bertani

LBS: LANA binding sites

LUR: Long unique coding region

LZ: Leucine zipper

MCD: Multicentric Castleman's disease

MHV-68: Murine gammaherpesvirus 68

miRNA: Micro RNA

MOI: Multiplicity of infection

MZ: Marginal zone

ncRNA: Non-coding RNA

NHEJ: Non-homologous end joining

NLS: Nuclear localization signal

O/N: Overnight

ORF: Open reading frame

Ori-lyt: Lytic DNA replication origin

Ori-P: Latent DNA replication origin

P: Proline-rich region

PAM: Protospacer adjacent motif

PCR: Polymerase chain reaction

PEL: Primary effusion lymphoma

PFU: Plaque forming unit

Q: Glutamine rich-region

RE: Replication element

RT: Room temperature

RTA: Replication and transcription activator

sgRNA: Single-guided RNA

TR: Terminal repeats

VZV: Varicella-zoster virus

WB: Western blot

WSM: Working stock media

WT: Wild-type

YFP: Yellow fluorescent protein

1. Introduction

1.1 Herpesviridae family

The order *Herpesvirales* is divided into three families, according to its host range: *Herpesviridae* (mammal, bird and reptile viruses), *Alloherpesviridae* (fish and amphibian viruses) and *Malacoherpesviridae* (mollusk viruses) (Davison *et al*., 2009).

Among *Herpesviridae* family (usually designated by herpesviruses) over 100 species have already been identified but it is likely that a lot more remain unidentified (Louten, 2016). Based on their biological properties (host range, replication cycle and cell tropism), this family is divided into three subfamilies: *alpha* (α)-, *beta* (β)- and *gamma* (γ)- *herpesvirinae* (Wen and Damania, 2010).

Herpesviruses are ubiquitous in nature infecting a wide range of hosts. Owing to their long co-evolution history with their respective hosts, viruses are well adapted to them and the infection is mostly asymptomatic, except in immunocompromised hosts and during viral transmission. Typically, they establish a long-term infection in their hosts which requires immune evasion (Davison *et al*., 2002; Grinde, 2013; Burrell *et al*., 2017). So far, eight herpesviruses have been associated with humans (Table 1.1.).

Table 1.1. Human herpesviruses and their key features (Adapted from Whitley, 1996; Davison *et al*., 2009).

Subfamily	Virus and respective abbreviation	ICTV[1] species name	Genome size (kbp)	Main site of latent infection
Alpha herpesvirinae	Herpes simplex virus type 1 (HSV1)	Human herpesvirus 1	≈152	Sensory nerve ganglia
	Herpes simplex virus type 2 (HSV2)	Human herpesvirus 2	≈154	
	Varicella-zoster virus (VZV)	Human herpesvirus 3	≈125	
Beta herpesvirinae	Human Cytomegalovirus (HCMV)	Human herpesvirus 5	≈230	Secretory glands, cells of the reticuloendothelial system and kidneys
	Human herpesvirus 6 (HHV-6)	Human herpesvirus 6	≈159	
	Human herpesvirus 7 (HHV-7)	Human herpesvirus 7	≈144	
Gamma herpesvirinae	Epstein-Barr virus (EBV)	Human herpesvirus 4	≈171	B cells
	Kaposi's Sarcoma-associated herpesvirus (KSHV)	Human herpesvirus 8	≈140	

[1] International Committee on Taxonomy of Viruses

These viruses are among the most complex in nature. Their virions contain double-stranded (ds) DNA located at the central core with a molecular weight higher than 100 kbp. This structure is surrounded by a capsid with icosapentahedral symmetry which is enclosed by a tight tegument. Surrounding this, there is a lipid bilayer envelope constituted by polyamines, lipids and glycoproteins. The latter ones are responsible for the different properties of each virus, providing unique antigens (Figure 1.1.). An interesting feature of these viruses is their genome since the number of open reading frames (ORFs) underestimates the genomic output due to alternative splicing and optional start codons (Whitley *et al*., 1996; Grinde, 2013).

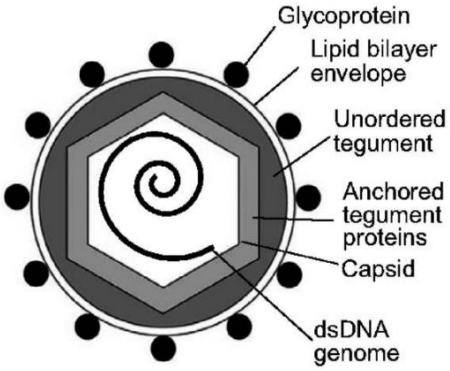


Figure 1.1. Schematic diagram of a human herpesvirus. The virion is constituted by dsDNA at the central core surrounded by an icosapentahedral capsid enclosed by a tegument which is enveloped by a lipid bilayer with polyamines, lipids and glycoproteins (Adapted from Liu and Zhou, 2007).

A characteristic feature of herpesviruses is their biphasic life cycle consisting of a lytic and a latent phase. The lytic phase is characterized by expression of viral ORFs in a temporally regulated cascade, leading to the production of new virions. In the latent phase, the viral genome is maintained in the nucleus of host cells as a circular, multicopy, extrachromosomal, non-integrated form (episome). Gene expression is highly restricted in this phase of infection, allowing the virus to evade the immune system. There are two main properties of latency: to persist, latent virus must be replicated and correctly segregated to daughter cells; and reversibility, characterized by the capacity to reactivate, in response to specific stimulus, entering the lytic phase to produce new infectious particles (Speck and Ganem, 2010).

1.2. *Gammaherpesvirinae* subfamily

The *Gammaherpesvirinae* subfamily has been identified in a range of animals from mice to man and infect more than 90% of humans (Nash *et al*., 2001; Godinho-Silva *et al*., 2014).

Based on DNA homology and genomic organization, this subfamily is divided into four genera: *Macavirus*, *Percavirus*, *Lymphocryptovirus* (gamma-1) and *Rhadinovirus* (gamma-2) (Simas and Efstathiou, 1998).

Epstein-Barr virus (EBV), a *Lymphocryptovirus*, and Kaposi's sarcoma-associated herpesvirus (KSHV), a *Rhadinovirus*, are the two gammaherpesviruses that have been so far identified in humans (Davison *et al*., 2009). These establish latency in the nucleus of lymphoid cell populations, particularly B cells, where the viral genome is maintained as an episome. They have also been related to human diseases, particularly, cancer (Wen and Damania, 2010).

1.3. Kaposi's sarcoma-associated herpesvirus (KSHV)

Kaposi's sarcoma (KS) was first described as an idiopathic multiple pigmented sarcoma of the skin in 1872 by Moritz Kaposi. However, only in 1994, Chang and Moore's group identified KSHV as

the infectious cause of KS (Chang *et al*., 1994; Wen and Damania, 2010). In 2010, the International Agency for Research on Cancer stated KSHV as a group 1 carcinogenic agent, demonstrating its public health importance (Rohner *et al*., 2014).

KS is an endothelial cell lineage tumor and the lesions can be found viscerally, cutaneously and mucosally. According to the epidemiology and clinical manifestations, it can be classified into four clinical subtypes: classic KS, endemic KS, iatrogenic/post-transplant KS and AIDS-associated KS (Cai *et al*., 2010; Wen and Damania, 2010). KSHV is necessary but not sufficient for KS development, being dependent of certain cofactors such as HIV infection (Mesri *et al*., 2010).

Besides KS, KSHV is also linked to two other disorders: primary effusion lymphoma (PEL), a non-Hodgkin's B cell lymphoma and plasmablastic variant of multicentric Castleman disease (MCD), a B cell lymphoproliferative disorder (Moore and Chang, 2003; Wen and Damania, 2010). It has also been associated with rare cases of bone marrow failure, hepatitis and acute inflammatory syndromes such as KSHV-associated inflammatory cytokine syndrome (KICS) and KSHV immune reconstitution inflammatory syndrome (IRIS) (Giffin and Damania, 2014; Mariggiò *et al*., 2017).

Seroepidemiological studies show that seroprevalence of KSHV infection is lower than 10% in northern Europe, Asia and United States, whereas in most of sub-Saharan Africa it is superior to 50%. The Mediterranean region presents an intermediate seroprevalence rate ranging from 10% to 30%. The transmission routes of KSHV are not clear but it has been proposed that both sexual and non-sexual modes are involved (Mesri *et al*., 2010; Rohner *et al*., 2014).

Nowadays, chemotherapy is one of the standard treatments for KS and PEL. Due to chemotherapy-relapse cases, molecular-targeted therapies are being developed and some drugs, such as imatinib, have already demonstrated to be a potential treatment option to multiresistant HIV-KS (Cao *et al*., 2015; Mui *et al*., 2017). However, the restricted host range of KSHV and subsequent absence of a robust small animal model constitutes a major limitation to investigate basic aspects of pathogenesis and treatments (Adler *et al*., 2000; Barton *et al*., 2011).

KSHV has a ds DNA genome consisting of a ≈ 140 kbp long unique region (LUR) which includes all KSHV ORF's, microRNAs (miRNAs), non-coding RNAs (ncRNAs) and antisense RNAs. This region is flanked by a variable number (≈40 to 50) of GC-rich 801 bp terminal repeats (TR) (Figure 1.2.) (Russo *et al*., 1996; Ohsaki and Ueda, 2012; Uppal *et al*., 2015).

To infect cells, the virion attaches itself to cell surface where it temporally interacts with cellular receptors. Since KSHV infects more than one cell type (such as endothelial cells, B cells, monocytes, epithelial cells, keratinocytes), it is likely that the virus interacts with ubiquitous cellular molecules (Chakraborty *et al*., 2012). After attaching to cells, the virion envelope fuses with plasma or endosome membrane allowing the entry of the capsid into the cytoplasm where it is dissembled, and the viral genome is released into the nucleus. There, using the cellular enzymatic machinery, the linear genome is circularized into an episome which is chromatinized to be maintained and to escape to the host DNA response mechanisms (Spear and Longnecker, 2003; Uppal *et al*., 2015).

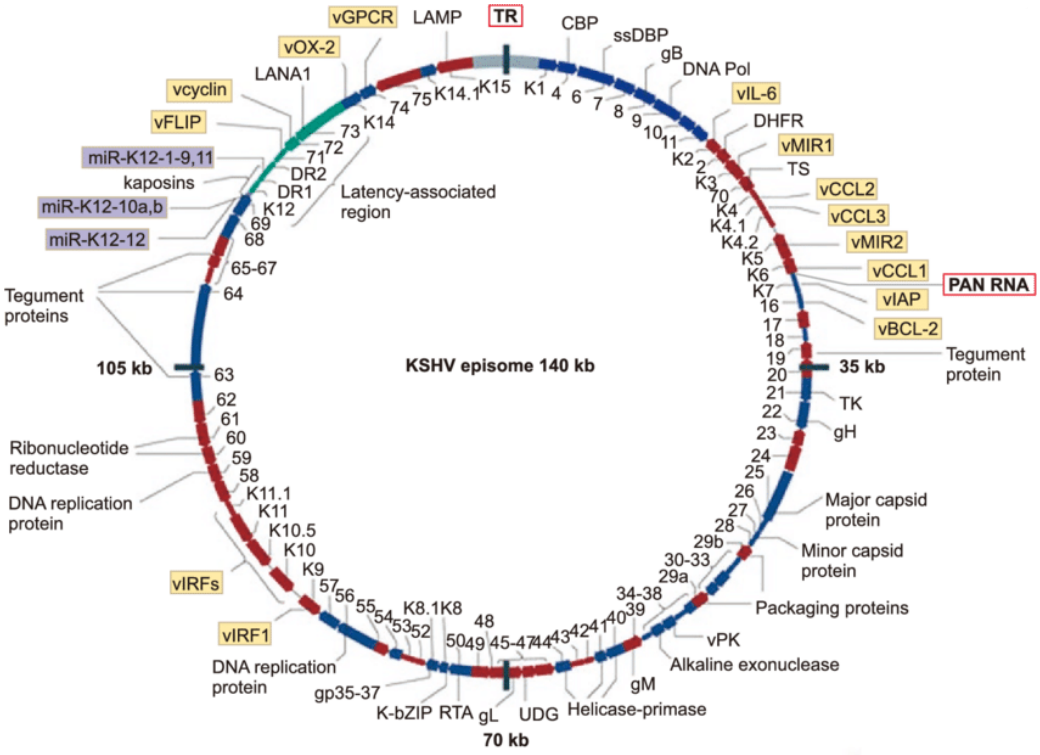


Figure 1.2. KSHV episome. It encodes 87 open reading frames and at least 17 miRNAs (purple boxes). The yellow boxes represent ORFs that encode cellular orthologues. The green boxes denote latent transcripts (Adapted from Avey *et al*., 2015).

Like all herpesviruses, inside the host cell nucleus, KSHV displays a life cycle constituted by a lytic and a latent phase (Figure 1.3.).

1.3.1. Lytic phase

This phase is characterized by expression of an ordered cascade of viral ORFs that will ensure viral linear DNA replication (on lytic origins of replication, ori-lyt) and its subsequent package into new viral particles which lead to cell lysis (Lin *et al*., 2003; Cai *et al*., 2010). There are three classes of lytic genes, according to the peak of their expression over time: immediate early (IE), early (E) and late (L) genes. While IE genes are involved in gene transcription, E genes encode proteins mainly related to DNA replication and gene expression, and L genes encode for structural and DNA packing proteins (Purushothaman *et al*., 2015). Besides these functions, some genes are also involved in blocking host's innate and adaptative recognitions (Kwun *et al*., 2007).

Therefore, this phase is indispensable for the preservation and dissemination of the virus and it also contributes to KSHV-induced tumorigenesis by inducing DNA double-strand breaks (DSBs) and impairing non-homologous end joining (NHEJ) repair mechanism (Xiao *et al*., 2013; Purushothaman *et al*., 2015).

Replication and transcription activator (RTA), an IE lytic master switch protein encoded by ORF50, is necessary and sufficient to initiate and drive lytic phase (Figure 1.3.) (Thakker and Verma, 2016).

1.3.2. Latent phase

Latency, a hallmark of gammaherpesviruses, is a non-productive phase characterized by an extra-chromosomal, circularized genome (episome) and a highly restricted gene expression with no viral particles being produced. The small subset of genes that are expressed in this phase are located in the major latency locus that includes four ORFs: ORF73 encoding for the latency-associated nuclear

antigen (LANA); ORF72 which encodes v-cyclin; ORF71 encoding for v-FLIP protein and ORFK12 which encodes kaposins A, B and C. In addition to these genes, there are also some miRNAs which are expressed (Figure 1.3.) (Cai *et al*., 2010). Even though the exact function of those miRNAs is unknown, they are likely to be involved in suppressing lytic reactivation and interfere with endothelial cell differentiation and angiogenesis. Since they do not encode proteins and cannot be recognized, the use of miRNAs may represent a smart viral mechanism to evade immune system (Coscoy, 2007; Mesri *et al*., 2010).

This phase allows the maintenance and long-term persistence of viral genome in the host cells while avoiding host immune responses due to viral gene expression restriction, which seems to be the major means to evade immune recognition during latency (Kwun *et al*., 2007). Besides gene expression restriction, there is also a downregulation of certain cell surface markers which are typically detected by the immune system (Cai *et al*., 2010). Latency is also essential for the pathogenesis of KSHV-associated malignancies by interfering with several cellular pathways (Mui *et al*., 2017).

Certain environmental and physiological factors may reactivate the virus to enter the lytic phase. These factors include inflammation, hypoxia, oxidative stress, epigenetic modifications, co-infection, host immune suppression, pharmacological agents (Cai *et al*., 2010; Purushothaman *et al*., 2015). The transition between latent and lytic phase is controlled by the interplay of RTA and LANA proteins which decide the cell fate.

The balance between lytic and latent phase is essential for the maintenance of the virus, since extending latent phase for a long time will cause the virus to vanish without production of viral progeny (Ohsaki and Ueda, 2012).

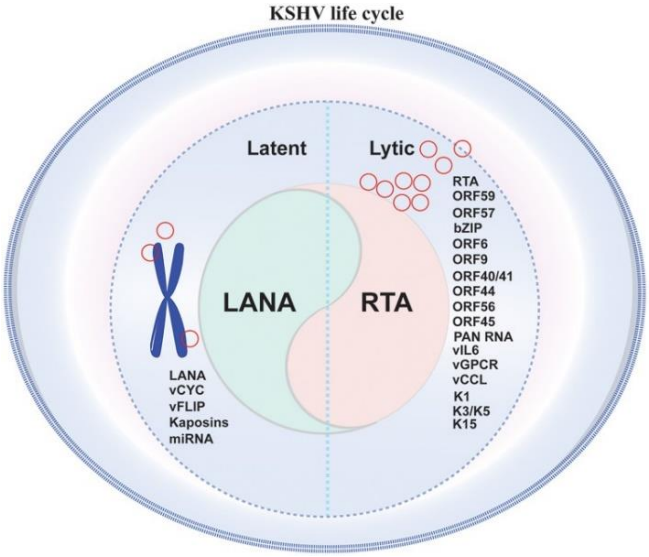


Figure 1.3. Representation of KSHV life cycle in an infected cell. Set of genes expressed in each phase are described (Adapted from Purushothaman *et al*., 2016).

1.3.3. kLANA

To be maintained, viral episomes must replicate with each cell division and segregate to daughter cells during mitosis.

KSHV LANA (kLANA) is a 1162 amino acid dimer protein encoded by ORF73. It is mostly nuclear, and it is one of the most important protein expressed during latent phase. This protein can be

divided into several regions, namely: a proline-rich N-terminal region, a glutamic acid-rich internal repeat central region and a C-terminal region (Uppal *et al*., 2014) (Figure 1.4.).

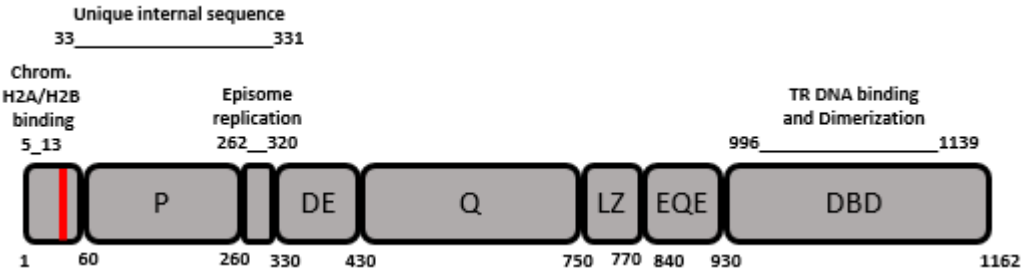


Figure 1.4. Representation of full-length kLANA protein. It is divided into several regions including a proline-rich region (P), an aspartate- and glutamate-rich region (DE), a glutamine-rich region (Q), a putative leucine zipper (LZ), a glutamate- and glutamine-rich region (EQE) and a DNA binding domain (DBD). The red bar denotes the nuclear localization signal (NLS), comprising amino acids 24 to 30. Amino acids 5 to 13 are involved in histones association. The region encompassing amino acids 33 to 331 (unique internal sequence) demonstrated to be important for segregation. For episome replication it has been demonstrated that amino acids 262 to 320 are crucial. Amino acids 996 to 1139 are required for the binding to terminal repeats (TR). Numbers indicate amino acids residues (Adapted from Piolot *et al*., 2001; De Leon Vázquez *et al*., 2013; Ponnusamy *et al*., 2015).

LANA is essential for the long-term persistence of viral episome and it is sufficient for the replication and persistence of TR-containing plasmids (Ballestas, 1999; Ye *et al*., 2004). Therefore, LANA plays an essential role in the segregation and replication of viral episome in proliferating cells. During latency, LANA acts on KSHV TR DNA to exert those functions. The N-terminal region interacts with histones H2A and H2B allowing the attachment of kLANA to mitotic chromosomes (Barbera *et al*., 2006; Uppal *et al*., 2014). The C-terminal region contains a DNA binding domain (DBD) that recognizes two LANA binding sites (LBS-1 and LBS-2) located in the TRs of the episome (Ballestas and Kaye, 2001; Garber *et al*., 2002; Hellert *et al*., 2015; Ponnusamy *et al*., 2015). By interacting with mitotic chromosomes and viral episome, LANA tethers the viral genome to host chromosomes allowing its segregation to daughter cells, being this function critical for viral persistence (Figure 1.5.).

Adjacent to LBS-1 and -2, there is a GC-rich 32 bp sequence called LBS-3 or replication element (RE). It is thought that the set LBS-RE constitutes the latent DNA origin of replication (ori-P). The binding of LANA to TR allows the initiation of viral DNA replication, probably by linking the ori-P and the host cellular replication machinery (Hu and Renne, 2005; Ohsaki and Ueda, 2012; Hellert *et al*., 2015). Recently, the region encompassing amino acids 262 to 320 demonstrated to be critical for DNA replication by recruiting replication factor C (RFC) to TR DNA (De Leon Vázquez *et al*., 2014; Sun *et al*., 2014; Li *et al*., 2015).

Recently, *in vitro* studies showed that the deletion of the region encompassing amino acids 33 to 273 resulted in reduced DNA replication. Host cell proteins involved in DNA replication have been described to interact with LANA, but none have been mapped to this region. Episome segregation was also assessed through green fluorescent protein (GFP) retention assay. Cells expressing LANAΔ33-273 were transfected with a GFP-based plasmid containing TR lacking RE. The mutant was highly deficient in the ability to retain GFP expression, suggesting an important role for the region encompassing amino acids 33 to 273 in the episome segregation. Deletions in the central repeat elements region indicated that this region might be dispensable for segregation (De León Vázquez *et al*., 2013; Juillard *et al*., 2016). The region 33 to 273 may interact with host cell proteins important for segregation. It is known that LANA interacts with SSRP1, a component of the FACT complex which is involved in the reorganization of nucleosomes. However, the protein has not yet been mapped to that region. Kinetochore protein Bub1 has been shown to interact with both N- and C-terminals. Disruption of Bub1

reduced the maintenance of KSHV episomes, supporting the hypothesis that Bub1 contributes to the persistence and segregation of KSHV genome (Xiao *et al*., 2010; Juillard *et al*., 2016).

More recently, our collaborators have mapped the amino acids 33 to 194 as being crucial to the segregation function, narrowing the region previously described (unpublished data).

LANA regulates both cellular and viral genes, interfering with several cellular signaling pathways to inhibit apoptosis and stimulate cell proliferation (Uppal *et al*., 2014; Mui *et al*., 2017). LANA also binds to viral promoters, such as RTA promoter, suppressing lytic gene transcription and controlling latency. This protein also associates with chromatin-modifying complexes (Weidner-Glunde *et al*., 2017).

In summary, LANA is a multifunctional protein essential for the persistence of viral episome and it also contributes to KSHV-induced tumorigenesis.

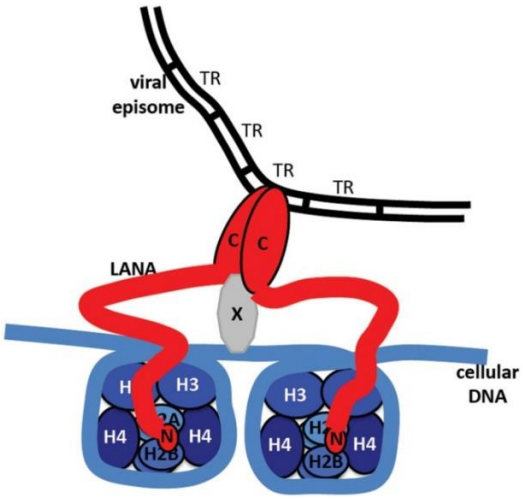


Figure 1.5. Mechanism of tethering used by kLANA. N-terminal region (N) interacts with histones H2A and H2B, and C-terminal region (C) associates with terminal repeats (TR) of viral episome. C-terminal region also interacts indirectly with mitotic chromosomes through a putative protein (X) (Adapted from Juillard *et al*., 2016).

1.4. Murine gammaherpesvirus 68 (MHV-68)

Murine gammaherpesvirus 68 is a gamma-2 herpesvirus that was first isolated from bank voles (*Clethironomys glareolus*) and yellow-necked wood mice (*Apodemus flavicollis*) (Blaskovic *et al*., 1980). Further studies suggested that MHV-68 is endemic in wood mice (*Apodemus sylvaticus)*, at least in United Kingdom, being a major reservoir for MHV-68 (Blasdell *et al*., 2003).

Analysis of gene organization among gammaherpesviruses showed a strong conservation of specific blocks of genes separated by virus-specific genes. This makes MHV-68, KSHV, Herpesvirus saimiri (HVS) and EBV genetically close viruses (Virgin *et al*., 1997).

The genome is composed by a 118 kbp region flanked by a variable number of 1,213 bp TRs. It encodes approximately 80 genes that are largely colinear with KSHV and HVS ORF's and some of them are homologous to cellular genes (Virgin *et al*., 1997).

Although it might not be its natural host, MHV-68 is capable of infecting laboratory mice (*Mus musculus*), contrarily to KSHV, establishing a chronic infection (Speck and Ganem, 2010). So, it

provides an animal model to study pathogenesis of gammaherpesviruses. Since there is a remarkable availability of mouse strains deficient for specific parameters, it is considered a major advantage over other models (such as primate models) (Adler *et al*., 2000). Besides that, genetic systems based on homologous recombination allow genetic engineering of the virus enabling the construction of mutant viruses (Speck and Ganem, 2010).

Following intranasal inoculation, lytic replication occurs transiently in lung alveolar epithelial cells, but it is cleared about two weeks post-infection. Dendritic cells (DCs) take the virus to lymph nodes and then it reaches the spleen by infecting marginal zone (MZ) macrophages first, and then MZ B cells. The latter ones go to the white pulp of the spleen, transferring the virus to follicular dendritic cells (FDC). Next, the virus goes to germinal center (GC) B cells, where a proliferation of latently infected B cells occurs, followed by the establishment of a long-term latency in memory B cells (Figure 1.6.) (Sunil-Chandra *et al*., 1992; Frederico *et al*., 2014). Other cells can also be latently infected, namely, macrophages, splenic dendritic cells and lung epithelial cells (Flaño *et al*., 2000; Speck and Ganem, 2010). The peak of latency is approximately 14 days post-infection which is coupled with a transient splenomegaly due to the proliferation of latently infected B cells, followed by the decrease of the expansion of infected cells to a steady-state level (Sunil-Chandra *et al*., 1992; Flaño *et al*., 2000).

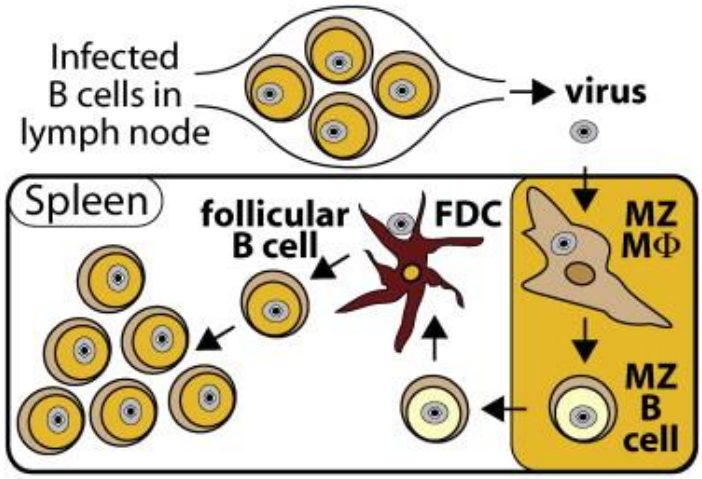


Figure 1.6. MHV-68 infection. The virus reaches the spleen by infecting marginal zone (MZ) macrophages and MZ B cells which are relocated to the white pulp of the spleen. Here the virus is transferred to follicular dendritic cells (FDC) and then reaches germinal center (GC) B cells (Adapted from Frederico *et al*., 2014).

1.4.1. mLANA

Like KSHV, MHV-68 ORF73 encodes a LANA nuclear protein (mLANA) which is homologous to kLANA, particularly the C-terminal region (amino acids 140 to 263) (Correia *et al*., 2013). However, as a result of the absence of the internal acidic and glutamine-rich repeat elements present in kLANA, mLANA is a smaller protein with 314 amino acids (Habison *et al*., 2012) (Figure 1.7.). This protein is expressed in germinal center B cells (Marques *et al*., 2003).

Studies concluded that mLANA mediate episome persistence since MHV-68 with mLANA disrupted was unable to establish an efficient latent infection in mice (Fowler *et al*., 2003; Moorman *et al*., 2003).

As kLANA, mLANA contains a DNA binding domain (DBD) (within C-terminal region) that acts on the TRs elements of the MHV-68 genome. Also, it was demonstrated that mLANA associates with mitotic chromosomes. Furthermore, both C- and N-terminals interact with cellular proteins (Habison *et al*., 2012; Correia *et al*., 2013).

Therefore, mLANA may has a similar tethering mechanism as kLANA, tethering MHV-68 episomes to mitotic chromosomes allowing a correct segregation to daughter cells (Habison *et al*., 2012).

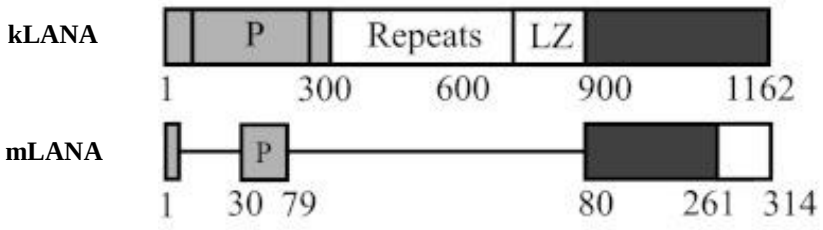


Figure 1.7. Comparison of full-length kLANA and mLANA proteins. Shaded regions represent homology between proteins. The intensity represents the level of homology (dark means more homology between regions). Unshaded regions lack homology. P: proline-rich region; LZ: Leucine zipper. Numbers indicate amino acids residues (Adapted from Habison *et al*., 2012).

1.5. Chimeric virus to study kLANA functions *in vivo*

A major limitation of the study of human gammaherpesviruses is the absence of a robust small animal model to investigate basic aspects of viral pathogenesis and therapeutics caused by the narrow host tropism of both EBV and KSHV.

Considering the gene conservation between MHV-68 and KSHV, the facility to introduce mutations and construct recombinant viruses and the capacity to infect laboratory mice, MHV-68 may represent a suitable model to study KSHV pathogenesis *in vivo*.

As described above, kLANA and mLANA share genetic and functional similarities. Habison and collaborators engineered a chimeric virus, designated by v.kLANA, in which the MHV-68 *mLANA* (v.WT) was replaced by kLANA ORF and its 5'UTR (Figure 1.8.). They demonstrated that kLANA is able to act on TRs from MHV-68 (mTR) and mLANA is also capable to act on TRs from KSHV (kTR). *In vivo* infection demonstrated that even though the latency of v.kLANA is lower than MHV-68 latency, the episome copy number of v.kLANA is similar to v.WT. Thus, v.kLANA is able to partially rescue MHV-68 latency *in vivo* (Habison *et al*., 2017).

In summary, kLANA and mLANA act reciprocally to maintain the viral episome providing a new model to study kLANA functions *in vivo*.

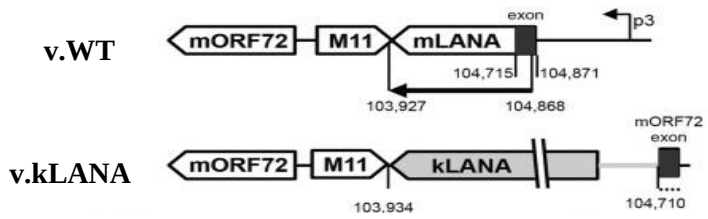


Figure 1.8. Schematic diagram of the construction of v.kLANA. p3 is a mLANA promoter (Adapted from Habison *et al*., 2017).

2. Aim of the project

KSHV is a gammaherpevirus that has a life cycle comprising a lytic and a latent phase. The latent phase is established in the B cells of the host and it is related with several malignancies. Even though there is a highly restricted gene expression, there are some genes expressed, namely, the kLANA coding gene. Studies show that this protein is critical for the replication and segregation of the virus, functions that will allow the maintenance of the virus. Therefore, a better understating of the functionality of this protein is essential for the development of strategies to control KSHV latent infection. However, KSHV has a narrow host, infecting only humans. This is a major limitation for *in vivo* studies.

MHV-68 is a gammaherpesvirus genetically related to KSHV that is able to infect laboratory mice (*Mus musculus*), establishing a persistent infection. It also expresses a LANA protein (mLANA) which is functionally homologous to kLANA. So, MHV-68 may offer a mouse model to study KSHV pathogenesis.

Previous work focused on the generation of a chimeric virus where *mLANA* is replaced by *kLANA* within MHV-68 genome. It was demonstrated that kLANA supports MHV-68 latent phase despite the lower levels observed. Thus, the chimeric virus may provide new insights into kLANA functions (Habison *et al*., 2017).

In vitro studies showed that the region encompassing amino acids 33 to 194 (of kLANA) has a negative impact on the segregation of the virus, and consequently on its maintenance (unpublished data).

The aim of this project was to study the effect of the deletion comprising amino acids 33 to 194 using the chimeric model and to assess the impact of the mutation *in vivo*. To accomplish this objective, a deletion between amino acids 33 to 194 of kLANA was generated. This mutated gene was introduced into MHV-68 genome replacing the original *mLANA*. The ultimate goal was to determine if the deletion of this region affected the establishment of latency in a mouse model of infection.

3. Materials and methods

3.1 Materials

3.1.1. Plasmids

The expression plasmid pSG5-T7-kLANAΔ33-194 (ampicillin (amp) resistance), which encodes kLANA containing a deletion between the amino acids 33 to 194 was provided by Kenneth Kaye (Harvard Medical School).

Plasmid pSP72_PCR1-5 (amp resistance), constructed by Marta Miranda (Habison *et al*., 2017), contains the full length kLANA ORF and its 5' UTR flanked by MHV-68 sequences.

pSP72_kLANAΔ33-194 was generated in this project by subcloning the region encompassing the 33-194 deletion from pSG5-T7-kLANAΔ33-194 into pSP72_PCR1-5 using *Bam*HI/*Sma*I sites. Details of cloning procedure are described in section 3.2.1.2.

BamG shuttle plasmid (kanamycin (kan) resistance) was generated by Sofia Marques. This plasmid contains MHV-68 genomic Bam-G fragment (coordinates 101,653 to 106,902) cloned into the

pST76K-SR (shuttle plasmid). The Bam-G fragment contains the mLANA coding sequence (coordinates 103,925 to 104,869).

BamG_kLANAΔ33-194 shuttle was generated in this project by subcloning the *Bgl*II fragment from pSP72_kLANAΔ33-194 into *Bgl*II sites of Bam-G shuttle, replacing mLANA by kLANA construct.

pGBT9-M9 plasmid is a pGBT9 based plasmid (amp resistance) containing MHV-68 M9 gene (coordinates 93,960 to 94,520), available at PSimas lab.

3.1.2. Bacteria

Escherichia coli (*E. coli*) XL 10-Gold Ultracompetent Cells (Agilent Technologies) were used to transform ligation reactions to obtain pSP72_kLANAΔ33-194 (section 3.2.1.2.1.).

The remaining ligation reactions from cloning procedures (section 3.2.1.2.2.) were transformed into competent *E. coli* DH5α made at PSimas lab.

E. coli DH10B, either containing MHV-68 bacterial artificial chromosome (BAC) (chloramphenicol (cam) resistance) (provided by Dr. Heiko Adler and Dr. Ulrich Koszinowski) or MHV-68 BAC.yfp (cam resistance) (provided by Dr. Samuel Speck) were used to generate recombinant viruses in non.yfp or yfp background, respectively.

3.1.3. Cell lines

Mouse embryonic fibroblasts expressing CRE recombinase (NIH3T3 CRE) were used to remove the BAC cassette. They were cultured in Dulbecco's modified emedium (DMEM) (Gibco® by life technologies) supplemented with 10% fetal bovine serum (FBS) (Gibco® by life technologies), 100 U/mL penicillin-streptomycin (Gibco® by life technologies) and 2 mM L-glutamine (Gibco® by life technologies).

Baby hamster kidney fibroblasts (BHK-21) were maintained in Glasgow minimum essential medium (GMEM) (Gibco® by life technologies) containing the same supplements as above and 10% tryptose phosphate broth (Sigma).

Both cell lines were maintained at 37ºC with 5% CO_2.

3.1.4. Viruses

MHV-68 (v.WT) was derived from a genomic BAC containing the MHV-68 genome, a gift from H. Adler and S. Efstathiou. This virus is MHV-68 clone G2.4 but contains a BAC vector (flanked by *lox*P sites) in its left end (Adler *et al*., 2000; Efstathiou *et al*., 1990).

MHV-68.yfp (v.WT.yfp) is a transgenic virus that expresses yellow fluorescent protein (YFP) driven by the human cytomegalovirus immediate-early promoter and enhancer, that is capable to mark infected cells during MHV-68 latency. The YFP expression cassette was introduced into the region between ORF27 and ORF29b (Collins *et al*., 2009). It was kindly provided by Dr Samuel Speck and was used in all yfp *in vitro* and *in vivo* assays as a control virus.

Chimeric kLANA MHV-68 virus (v.kLANA), in which *mLANA* was replaced by *kLANA* and its 5'UTR, expressing or not YFP, was constructed by Marta Miranda, as described in Habison *et al* (2017). It was used in all experiments as a control.

3.1.5. Mice

Six to 8-week-old female C57BL/6J mice were purchased from Charles Rivers Laboratories, France. Mice were housed and subjected to experimental procedures in specific pathogen free conditions, at Instituto de Medicina Molecular João Lobo Antunes rodent facility, Lisbon, Portugal.

3.2. Methods

3.2.1 Construction of chimeric virus MHV-68-kLANAΔ33-194

3.2.1.1 Isolation and analysis of DNA

3.2.1.1.1. Restriction endonuclease reactions

Restriction endonuclease reactions were performed to obtain inserts, linear vectors and to screen clones. All reactions were performed using 500-4000 ng of DNA and 1-20 U of restriction enzyme (New England Biolabs or Roche) in a total volume of 30-50 µL and incubated for 1 hour (h) to overnight (O/N), at the temperature recommended by the enzyme manufacturer.

3.2.1.1.2. Polymerase chain reaction (PCR)

PCR was performed to amplify fragments for cloning procedure or to confirm the presence of the deletion. The reaction mix contained 1× Pfu DNA Polymerase buffer with $MgSO_4$, 200 µM of each dNTP, 300 nM of each primer, 50 ng of DNA, 1.25 U Pfu DNA polymerase (Promega) and ultrapure water (MiliQ) up to 50 µL. DNA was amplified on a thermocycler (Biorad) with the following parameters: initial denaturation at 95ºC for 2 minutes (min); 30 cycles comprising 3 phases – denaturation at 94ºC for 45 s, annealing at 50ºC for 30 s and extension at 72ºC for 2 min/kbp. A final step of extension was done at 72ºC for 5 min.

3.2.1.1.3. Analysis and isolation of DNA by gel electrophoresis

DNA fragments were separated by size on 0.8%-1.5% agarose gels (NZYtech), in 1×TAE (40 mM Tris-acetate, 1 mM EDTA, pH 8.0) and DNA was stained with red safe (iNtRON Biotechnology) or gel red (Biotium). Loading buffer (10 mM EDTA, 5% glycerol, 0.025% bromophenol blue, 0.025% xylene cyanol) was mixed with the samples before loading them into the gel. Electrophoresis was performed 1 h to O/N, at 50 V to 90 V, depending on the assay. DNA fragments were visualized by UV transillumination and analyzed by comparison with the molecular weight ladder (1 kb DNA plus, Invitrogen). If applicable, the band of interest was excised from gel and DNA was purified using QIAquick gel extraction kit (Qiagen), following the manufacturer's instructions. A small amount of purified DNA was analyzed on agarose gel to check DNA purification.

3.2.1.1.4. DNA Ligation

Vectors and inserts digested and purified were ligated using T4 DNA ligase (NEB). A molar ratio of 1:3 vector to insert were ligated in a 20 µL reaction containing 2 µL 10× T4 DNA ligase buffer, 1 U T4 DNA ligase and MiliQ water. The reactions were incubated O/N at 16ºC.

3.2.1.1.5. DNA quantification and sequencing

DNA was quantified by ultraviolet spectrophotometry at 260 nm using a ND-2000 spectrophotometer (Thermo Scientific). To confirm the sequence of *kLANA* encompassing the deletion, DNA plasmid minipreps were sequenced (Supplementary data, table S1) at GATC (Ebersberg, Germany) using the Sanger method. The sequences were analyzed with SnapGene 4.0.7 software.

3.2.1.1.6 DNA plasmid isolation – Miniprep

In order to isolate plasmid from *E. coli*, a single colony was inoculated in 10 mL of Luria-Bertani (LB) broth (NZYtech) supplemented with the appropriate antibiotic (100 µg/mL amp and/or 30 µg/mL kan) and incubated overnight at 37ºC or 30ºC (depending on the plasmid). Cultures were centrifuged at 3059×*g* for 10 min, and the pellet was processed using Wizard Plus SV minipreps DNA

Purification System (Promega), accordingly to manufacturer's instructions. Finally, DNA was eluted in 70 µL of nuclease free water and stored at -20 ºC.

3.2.1.1.7. BAC DNA preps

BAC-minipreps were obtained by using the alkaline lysis method followed by phenol/chloroform extraction. A bacterial culture grown O/N at 37 ºC and 220 rpm was centrifuged for 10 min at 3059×*g* and at 4ºC. After discarding the supernatant, the pellet was evenly resuspended in 200 µL of buffer S1 (NucleoBond BAC 100 kit, Macherey-Nagel™) and 200 µL of buffer S2 (NucleoBond BAC 100 kit) were added to lyse cells. The resulting suspension was carefully mixed by inversion and incubated for 5 min, at room temperature (RT). To neutralize the suspension, 200 µL of pre-chilled buffer S3 (NucleoBond BAC 100 kit) were added, followed by gentle mixing and a 15 min incubation step on ice. To separate cell debris from DNA, the suspension was centrifuged for 15 min at 17 949×*g* at 4ºC. The supernatant (500 µL) was transferred to a new tube and the same volume of phenol:chloroform:isoamyl alcohol (25:24:1, Applichem) was added. After mixing by inversion and an incubation of 5 min at RT, a step of centrifugation at 17 949×*g*, at 4ºC for 8 min was done. The aqueous phase (400 µL) was carefully transferred, with a cut tip, to a new tube, and 0.7 volumes of isopropanol (Sigma) were added. The suspension was mixed and incubated 5 min at RT. BAC DNA was pelleted by centrifugation for 20 min at 17 949×*g* at 4ºC. The supernatant was discarded and 200 µL of 70% ethanol were added to wash the pellet. A last step of centrifugation was performed at 17 949×*g* for 10 min at 4ºC. After air drying, the pellet was resuspended in 60 µL of nuclease-free water and stored at 4ºC.

Since large amounts of pure BAC DNA are necessary for transfections, BAC DNA maxipreps were obtained. 300 mL of bacterial culture were grown O/N at 37ºC and 220 rpm. The purification was done using NucleoBond BAC 100 kit (Macherey-Nagel™), according to manufacturer's instruction.

3.2.1.2. Cloning procedures

3.2.1.2.1. Subcloning of kLANAΔ33-194 into pSP72_PCR1-5

The In-fusion HD cloning kit (Takara) was used to clone a DNA fragment encoding the kLANAΔ33-194 into pSP72_PCR1-5**.** Firstly, a PCR reaction (section 3.2.1.1.2.) using primers IMM_NS2 and IMM_NS3 (Supplementary data, Table S1) amplified the region between *Bam*HI and *Sma*I of pSG5-T7-kLANAΔ33-194 which contains kLANAΔ33-194 coding DNA. The PCR product and *Bam*HI/*Sma*I - cut pSP72_PCR1-5 were gel extracted (section 3.2.1.1.3.). Purified fragments were ligated using the In-Fusion kit according to the manufacturer's instructions. Ligations were transformed into *E. coli* XL-10 Gold ultracompetent cells (section 3.2.1.3.1.). Plasmid was isolated from colonies by plasmid miniprep (section 3.2.1.1.6.) and the region encompassing the deletion was sequenced (section 3.2.1.1.5.). After analyzing the results, two errors (point mutations) were detected. To correct those errors and to remove a restriction site for *Bgl*II, introduced by the deletion, QuikChange Multi Site-Directed Mutagenesis Kit (Agilent technologies) was used with primers IMM_NS3 and IMM_NS4 (Supplementary data, Table S1), following manufacturer's instructions. The reaction was transformed into *E. coli* XL-10 Gold Ultracompetent cells (section 3.2.1.3.1.). Correction of the errors and removal of the *Bgl*II site were confirmed by sequencing. The resulting plasmid was named pSP72_kLANAΔ33-194 (Supplementary data, Figure S1).

3.2.1.2.2. Subcloning of pSP72_kLANAΔ33-194 insert into BamHI-G shuttle vector

The BamG shuttle and pSP72_kLANAΔ33-194 were digested with *Bgl*II (Roche) and insert and vector bands were purified after gel excision (section 3.2.1.1.3.). Ligation between insert and vector was performed and the resulting reaction was transformed into DH5α cells (section 3.2.1.3.1.). Plasmid DNA was isolated from individual colonies by DNA miniprep and screened using the appropriate

endonucleases. The resulting plasmid was named BamG_kLANAΔ33-194 shuttle (Supplementary data, figure S1).

3.2.1.3. Bacterial methods

3.2.1.3.1. Transformation of competent cells

E. coli DH5α competent cells, *E. coli* XL 10-Gold ultracompetent cells and *E. coli* DH10B were transformed using the heat shock method. After gently thawing an aliquot on ice, an amount of DNA was added to the cells. The tubes were incubated on ice for 30 minutes. Cells were heat-shocked for 30 seconds (*E. coli* XL-10 Gold) or 90 seconds (*E. coli* DH5α and *E. coli* DH10B). 0.5 mL of Super Optimal Broth (SOC) (NZYtech) were added and cells were incubated for 1 h at 37ºC or 30ºC (depending on the plasmid) with agitation of 220 rpm. Cells were spread into LB agar (NZYtech) plates containing the appropriate antibiotic (100 µg/mL ampicillin, 30 µg/mL kan and/or 17 µg/mL cam) and incubated overnight at 37ºC or 30ºC.

3.2.1.3.2. Mutagenesis

To generate MHV-68 recombinant viruses, mutagenesis was performed according to the Two-Step-Replacement strategy (O'Connor *et al*., 1989). Firstly, BamG_kLANAΔ33-194 shuttle was transformed into *E. coli* DH10B competent cells harboring a BAC with the wild-type (WT) MHV-68 genome (non.yfp and yfp background). Bacteria were spread in LB plates containing kan and cam and were incubated 1½-2 days at 30 ºC. In this step, a homologous recombination event between MHV-68 sequences in the shuttle BamG_kLANAΔ33-194 and MHV-68 BAC occurred, mediated by protein RecA which is encoded by shuttle plasmid, enabling the integration of BamG_kLANAΔ33-194 shuttle plasmid into the BAC genome, forming co-integrates.

To select co-integrates, 10 single colonies (from each background) were spread in LB plates containing kan and cam and were incubated overnight at 43 ºC. To make sure that the clones are clonal, this last step was repeated. Next, bacteria were spread into cam plates and incubated 1½ days at 30 ºC. Here, co-integrates can resolve themselves by spontaneous homologous recombination to their initial condition or to BAC mutant state.

To select clones with the resolved BAC mutants, 10 individual colonies (from each background) were re-streaked into cam plus 5% sucrose (Sigma) plates, which is a counter selection against *sacB* gene presented in shuttle plasmid and were incubated for 1½ days at 30 ºC.

From each of the 10 plates (from each background), 5 clones were selected and streaked on kan plates and in parallel on cam plates and incubated O/N at 37ºC. Kan-sensitive and cam-resistant clones indicate that co-integrates are well resolved.

Analysis of those clones was performed by colony PCR using IMM_NS2 and IMM_NS3 (Supplementary data, Table S1) to confirm the presence of the *kLANAΔ33-194* mutant. From those clones that were positive, BAC plasmids were extracted by BAC miniprep (section 3.2.1.1.7.) and screened by enzymatic digestion with *Eco*RI (Roche). Clones with the expected restriction profile were selected to grow to prepare BAC maxipreps. To assess the restriction profile of the selected clones, they were digested with *Eco*RI, *Hind*III (Roche) and *Bam*HI (Roche).

3.2.2. Reconstitution of MHV-68_kLANAΔ33-194

3.2.2.1 Virus Reconstitution on BHK-21

BHK-21 cells were seeded one day before transfection at 1×10^6 cells/6 cm^2 dish. One µg of BAC DNA maxiprep was added, with cut tip, to 500 µL of unsupplemented GMEM in a microtube. Two µL of FuGENE® HD transfection reagent (Promega) were added into the media. After gently

mixing, the transfection mix was incubated for 20 min, at RT, and then added to the cells. Cells were incubated at 37ºC for 3½ days until cytopathic effect (CPE) was observed. Cells were scrapped onto media, harvested and stored at -80 ºC. These were designated by BAC+ stock virus.

3.2.2.2. Passage of virus through NIH3T3-Cre cells to remove BAC sequences

Since BAC cassette (which is flanked by *lox*P sites) contains a *gfp* gene (Adler *et al*., 2000; Adler *et al*., 2001), the screening of virus containing those sequences can be performed using GFP as a marker. BAC+ viruses were passed through NIH3T3-Cre cells to remove sequences between *lox*P sites. One day before infection, 1.75×10^5 cells/well were seeded in 6-well plates. Limiting dilutions of BAC+ were done and within few cell passages GFP signal was not observed, using a fluorescent microscope. Cells were scrapped, harvested and stored at – 80ºC. These viral stocks were designated by BAC- stocks and were used to produce working stocks.

3.2.3. *In vitro* assays

3.2.3.1. Production of viral stocks

Working viral stocks were produced from BAC- viruses, by infecting 5×10^6 BHK-21 cells at a low multiplicity of infection (MOI) (0.002 PFU/cell). After 4 days of incubation at 37ºC, cells were scrapped onto media and centrifuged at 1 500×*g* for 5 min at 4 ºC. The cell pellet was resuspended in 2 mL of GMEM, aliquoted and stored at -80ºC – Cell Working Stocks (CWS). The supernatant was centrifuged at 15 000×*g* for 2 h at 4ºC. Pellet was resuspended in 2 mL of GMEM, aliquoted and stored at -80ºC – Working Stock Media (WSM). Those were used in all infectivity assays.

3.2.3.2. Virus titration using suspension assay – plaque assay

BAC- and WSM were titred by suspension assay. For this, 900 µL of GMEM were distributed to 15 mL falcon tubes, 100 µL of virus were added to the first falcon and 10-fold serial dilutions were made. Next, 1 mL of 2×10^5 cells was added to each falcon and tubes were incubated in a rotating table (30 rpm) at RT for 1 h. Thereafter, 2 mL of GMEM were added to the tubes, which were mixed by inversion and the suspension was added to 6-well plates. Following an incubation of 4 days at 37 ºC, media was removed, and cells were fixed with 4% formaldehyde (in PBS), for 10 min. Cells were stained with 0.1% toluidine blue, for 5 min. Viral plaques were analyzed using a magnifying glass and titers were calculated accordingly to the following formula:

$$Virus\ titer = No.\ of\ plaques \times \frac{1}{Dilution} \times \frac{1}{Inocolum}$$

3.2.3.3. Multi-step growth curve

Multi-step growth curves were performed in 24-well plates. 5×10^4 BHK-21 cells/well were infected with a MOI of 0.01 PFU/cell in 200 µL of GMEM and were incubated at 37ºC for 1 h. After the incubation, inoculum was removed, and cells were washed twice with 500 µL of PBS. One mL of GMEM was added to each well and cells were incubated at 37ºC. Cells were scraped onto media and harvested on the following time points: 0 h (right after addition of GMEM), 24 h, 48 h, 72 h, 96 h and 120 h after infection. Samples were kept at -80ºC until titration using suspension assay method (section 3.2.3.2.).

3.2.3.4. Immunoblotting

BHK-21 cells (5×10^5 cells/well of a 12-well plate) were infected with v.WT, v.WT.yfp, v.kLANA, v.kLANA.yfp, v.kLANAΔ33-194 and v.kLANAΔ33-194.yfp with a MOI of 3 PFU/cell in 500 µL of GMEM. A negative control well was inoculated with 500 µL of supplemented GMEM. Cells were incubated for 2 h at 37 ºC with gentle rocking every 30 min. Next, the inoculum was removed,

and the cells were washed with 1 mL of pre-warmed GMEM. One mL of media was added to each well, followed by an incubation period of 4 h at 37 ºC. After, cells were washed twice with ice-cold PBS and 110 µL of lysis buffer (10 mM Tris-HCl pH 7.4, 150 mM NaCl, 1 mM Na_3VO_4, 1 mM NaF, 1% Triton™ X-100, protease inhibitors (Roche), miliQ H_2O) were added. Following the harvesting of cells into a tube, they were incubated for 10 min on ice and centrifuged for 10 min at 17 949×*g* at 4 ºC. Supernatants (100 µL) were transferred into a new tube and 100 µL of 2× Laemmli's buffer (100 mM Tris-HCl pH 6.8, 20% glycerol, 4% SDS, 10% β-mercaptoethanol) were added. Protein samples were heated at 100 ºC for 3 min and 50 µL from each one were separated by SDS-PAGE, in 1.5 mm mini-slab gels (BioRad). Five µL of molecular weight marker (Dual color, BioRad) were also loaded into the gel. A 10% resolving gel was prepared with 30% polyacrylamide (Acrylamide:bis-acrylamide 37.5:1, BioRad), 1.5 mM Tris-HCl (pH 8.8), 10% SDS (BioRad), 10% ammonium persulfate and tetramethylethylenediamine (TEMED) (Sigma). Stacking gel was prepared with 30% polyacrylamide, 1.0 M Tris-HCl (pH 6.8), 10% SDS, 10% ammonium persulfate and TEMED. Electrophoresis was performed at 180 V in 1×TGS buffer (25 mM Tris, 192 mM glycine and 0.1% SDS) (Biorad). After separation, proteins were transferred to a nitrocellulose membrane (GE Healthcare) by wet-transfer method, using 1×TG buffer (25 mM Tris, 192 mM glycine, 20% methanol) (BioRad) and 0.1% SDS, O/N at 100 mA and 4ºC. Protein transference was confirmed by incubation of the membrane with Ponceau S dye (Sigma) for 2 min. The dye was removed by incubation with washing buffer (PBS+0.05% Tween-20). To reduce non-specific background, the membrane was incubated with blocking solution (5% skimmed milk powder in washing buffer) on a rotating platform for 30 min. After washing the membranes, they were incubated with primary antibodies diluted in blocking buffer (Supplementary data, Table S2) for 1 h to O/N on a rotating platform. Membranes were washed and incubated with horse radish peroxidase (HRP)-conjugated secondary antibodies (Supplementary data, Table S3) for 30 min-1h on a rotating platform. Membranes were washed 3 times for 5 minutes. Detection of protein-antibody complexes were performed by chemiluminescence using SuperSignal® West Pico Chemiluminescent Substrate (Thermo Scientific), according to manufacturer's instructions.

3.2.4. *In vivo* assays

3.2.4.1 Ethics statement

This study was performed in accordance with the Portuguese Official Veterinary Directorate, European Directive 2010/63/EU and Federation of European Laboratory Animal Science Associations guidelines on laboratory animal welfare.

3.2.4.2. Infection of mice

Six- to eight-week-old mice were anaesthetized with isoflurane and were intranasally inoculated with 10^4 PFU of v.WT.yfp, v.kLANA.yfp and recombinant viruses in 20 µL PBS. Fourteen days post-infection, mice were sacrificed by cervical dislocation and spleens were collected to a falcon containing 5 mL of GMEM.

3.2.4.3. Single cell suspension

Spleens were homogenized and filtered through a 100 µm cell strainer to remove cell debris. After centrifuging cells at 275×*g* for 10 min at 4ºC, pellet was resuspended in 1 mL of Red Blood Lysis Buffer (154 mM ammonium chloride, 14 mM sodium hydrogenate carbonate, 1 mM EDTA pH 7.3) and incubated on ice for 5 min. Five mL of GMEM were added and a step of centrifugation was performed at 275×*g* for 5 min at 4ºC. Pellet was resuspended in 1 mL of GMEM and then 4 mL of GMEM were added to a final volume of 5 mL. This final suspension (single cell suspension) was used to assess the viral latency by infectious center assay (ICAs), the frequency of viral DNA positive cells by limiting dilution assay and the frequency of infected GC B cells by flow cytometry analysis.

3.2.4.4. Infectious center assay (ICAs)

Latent virus titers were determined by reactivation assay. Two-fold and 10-fold serial dilutions of splenocyte suspension were prepared, in duplicate, and 1 mL was added to 6 cm^2 dishes containing 1 mL of 4×10^5 BHK-21 cells and media up to 5 ml. Plates were incubated for 5 days at 37ºC. Spleens were also screened for the presence of pre-formed infectious viruses. The splenocyte suspension was freeze-thawed, plated and incubated for 4 days at 37ºC.

In both assays, cells were fixed with 4% formaldehyde in PBS and stained with 0.1% toluidine blue. Viral plaques were analyzed using a magnifying glass and titers were calculated according to the formula previously described (section 3.2.3.2.).

3.2.4.5. Limiting dilution assay and analysis by real-time PCR

To analyze the frequency of viral genome positive cells, a splenocyte pool for each infectious group was prepared from single cell suspensions from individual mice. The pool was centrifuged at 179×*g* for 5 min at 4ºC and pellet was resuspended in 2% FBS/PBS. This step was repeated twice. After the last resuspension, the sample was filtered using a 40 µm cell strainer. Next, cells were counted to prepare a dilution of 2×10^6 cells in 100 µL 2% FBS/PBS and serial 2-fold dilutions were done in 2% FBS/PBS.

Five µL of each of the eight replicates from each dilution were pipetted into PCR tubes containing 10 µL of lysis buffer (10 mM Tris-HCl, 3 mM, $MgCl_2$, 50 mM KCl, 0.45% NP-40, 0.45% Tween-20, 0.5 mg/mL Proteinase K). PCR tubes were incubated O/N at 37ºC. Proteinase K was inactivated by heating cell lysates to 95ºC for 5 min. Lysates were kept at -20ºC until further use.

Samples were analyzed by real-time PCR on Rotor Gene 6000 (Corbett Life Science) using a fluorescent Taqman probe and primers specific for MHV-68 M9 gene. The PCR mix contained 200 nM of each primer (Supplementary data, Table S1), 1× mix (Invitrogen), 300 nM of probe M9-T (6-FAM-CTTCTGTTGATCTTCC-MGB) (Alfagene), 5 mM of $MgCl_2$ and nuclease-free water up to 22.5 µL. To 22.5 µL of PCR mix, 2.5 µL of each sample were added to each tube. A negative control (water) and a positive control (pGBT9-M9) were analyzed together with the samples. The program had the following settings: first hold for 2 min at 50ºC followed by a second hold for 10 min, at 95ºC. The cycling step was repeated 40 times and consisted in a first phase of 15 sec at 95ºC and a second phase for 30 s at 60ºC. The acquisition was performed using the green channel with a gain value of 5.33. Results were analyzed using Rotor Gene 6000 software. Each sample was scored positive or negative comparing to negative and positive controls results. The frequency of viral DNA positive cells was calculated according to the single-hit Poisson model (SHPM) by maximum likelihood estimation (Bonnefoix *et al*., 2001). The model assumes that one cell of one cell subset is necessary and sufficient for generating a positive response. The method developed by Bonnefoix *et al* (2001) consists in modelling the limiting dilution data according to the linear log-log regression model.

3.2.4.6. Flow cytometry

One mL of individual single cell suspensions (section 3.2.4.3.) was centrifuged for 5 min at 275×*g*. The pellet was resuspended in 1 mL of 2% FBS/PBS and filtered through a 40 µm cell strainer. Splenocytes from each sample were counted and 10×10^6 cells were seeded in a round-bottom 96-well plate. Then, cells were incubated for 15 min at 4ºC with Fc blocking solution (Supplementary data, Table S4) diluted in 2% FBS/PBS. Cells were washed with 100 µL of 2% FBS/PBS and pelleted by centrifugation for 1 min at 765×*g* at 4ºC. Splenocytes were stained with the appropriate antibodies (Supplementary data, Table S4) diluted in 2% FBS/PBS, for 25 min at 4ºC in the dark. Cells were washed twice with 2% FBS/PBS to remove unbound antibodies. After resuspending cells in 2% FBS/PBS, they were transferred to FACS tubes that were kept on ice and protected from light until

analysis. Samples were analyzed on LSR Fortessa flow cytometer (BD Biosciences) using FACSDiva software for acquisition and FlowJo for analysis.

3.2.4.7. Statistical analysis

Statistical analysis was performed by Mann-Whitney test or one-way ANOVA, using GraphPad Prism software. For limiting dilution assay, 95% confidence intervals were determined as described by Marques *et al* (2003).

4. Results

4.1. Generation and characterization of MHV-68 recombinant virus

Recently, *in vitro* studies showed that the region between amino acids 33 to 194 of kLANA was required for segregation of viral episomes through mitosis (unpublished data). To evaluate the importance of the region *in vivo* using the chimeric model, a recombinant virus was constructed with a deletion encompassing amino acids 33 to 194 within *kLANA* which was introduced into the MHV-68 genome, replacing endogenous *mLANA* (Figure 4.1.). This virus was generated in two MHV-68 genomic backgrounds: yfp background (v.kLANAΔ33-194.yfp), where the recombinant virus expresses the YFP protein enabling the tracking of infection through flow cytometry analysis; and non.yfp background (v.kLANAΔ33-194) that does not express the YFP protein.

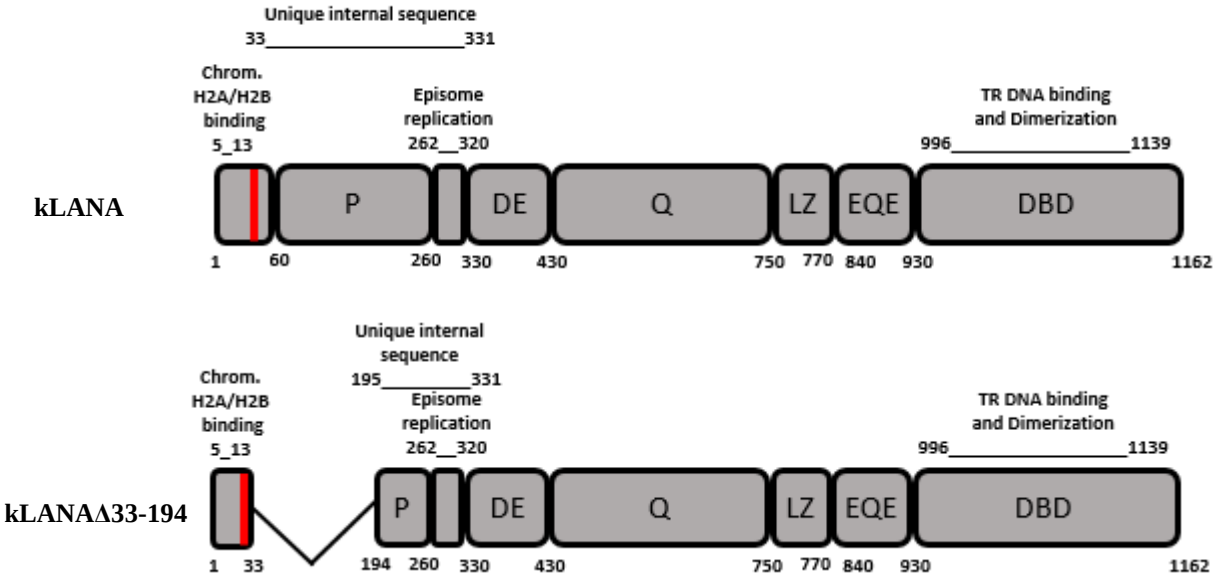


Figure 4.1. Schematic representation of wild-type kLANA and mutant kLANA (kLANAΔ33-194) proteins. The protein domains are indicated: proline-rich region (P), aspartate and glutamate (DE), glutamine (Q), leucine zipper (LZ), glutamate and glutamine (EQE) and DNA binding domain (DBD). The red bar represents the nuclear localization signal (NLS). kLANAΔ33-194 contains a deletion between amino acids 33 to 194. The numbers indicate amino acids residues.

The generation of the recombinant viruses was performed through a mutagenesis procedure in *E. coli* DH10B harboring a BAC plasmid with the MHV-68 genome cloned (MHV-68 BAC), according to the Two-Step-Replacement-Strategy (O'Connor *et al*., 1989). In this strategy, a process of homologous recombination occurs between the BamG_kLANAΔ33-194 shuttle plasmid and the MHV-

68 BAC allowing the introduction of mutant *kLANA* and it's 5'UTR into MHV-68 genome, replacing mLANA coding sequence (Adler *et al*., 2000).

To identify BACs containing mutant *kLANA* instead of *mLANA*, a colony PCR with *kLANA* specific primers for the region encompassing the deletion was performed. Clones that presented the correct band were grown and BAC DNA was extracted and analyzed by restriction digestion with *Eco*RI. From those which had the correct restriction profile, two clones (one from each background) were selected to prepare high quality maxiprep BAC DNA, and genome integrity was assessed by restriction enzyme digestion using 3 endonucleases: *Bam*HI (Figure 4.2. A.I and B.I), *Eco*RI (Figure 4.2. A.II and B.II) and *Hind*III (Figure 4.2. A.III and B.III), taking v.WT, v.WT.yfp ,v.kLANA and v.kLANA.yfp as controls. Since the observed profiles were according to the expected (Figure 4.3. I and II), mutant BAC DNA was transfected into BHK-21 cells to reconstitute viral progeny. BAC cassette is flanked by *lox*P sites, therefore BAC sequences were excised through Cre-lox system by infecting NIH-3T3 cells expressing Cre recombinase with the viruses.

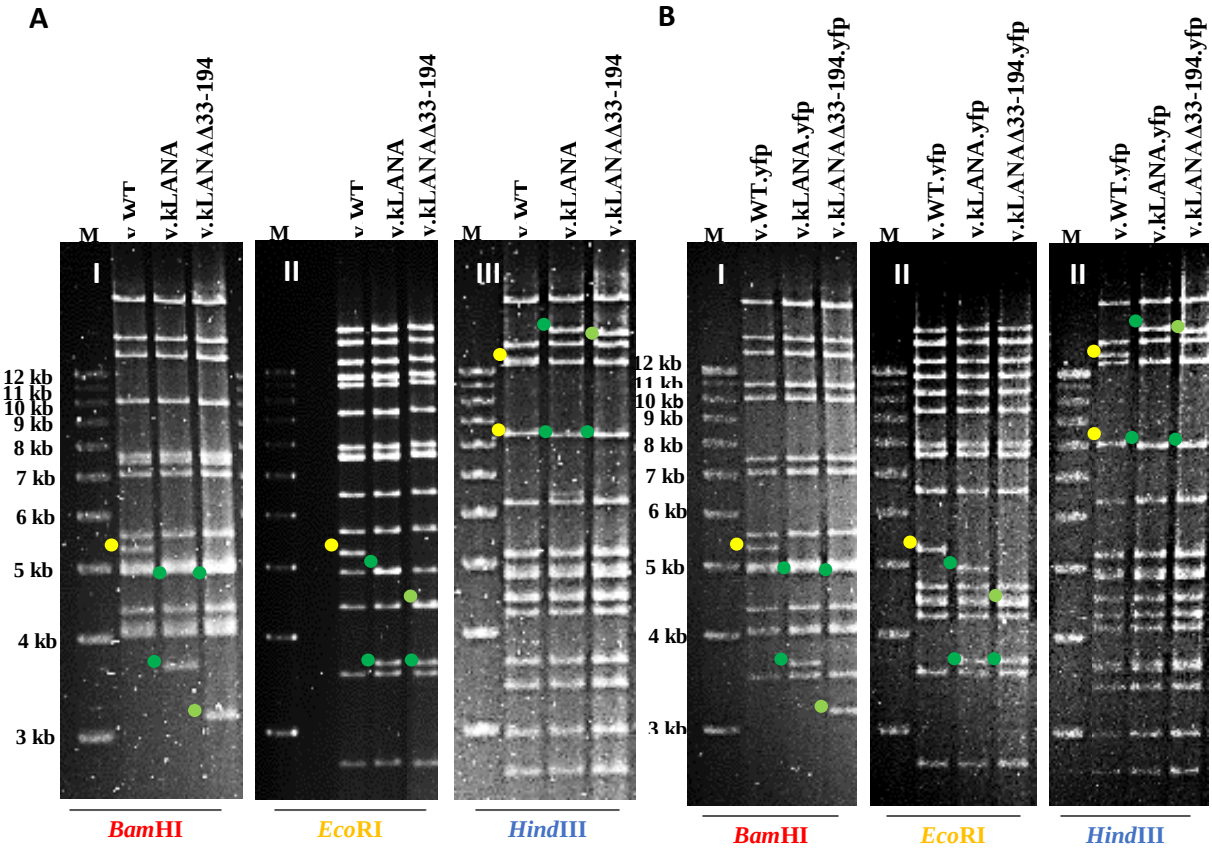


Figure 4.2. Restriction profiles of BAC plasmids (v.WT, v.kLANA and v.kLANAΔ33-194), in two different backgrounds: non.yfp (A) and yfp (B), to assess genome's integrity. I) Restriction profile using *Bam*HI; II) Restriction profile using *Eco*RI; III) Restriction profile using *Hind*III. Yellow, dark green and light green dots indicate specific bands of v.WT, v.kLANA and v. kLANAΔ33-194, respectively. Molecular weight is indicated on the left.

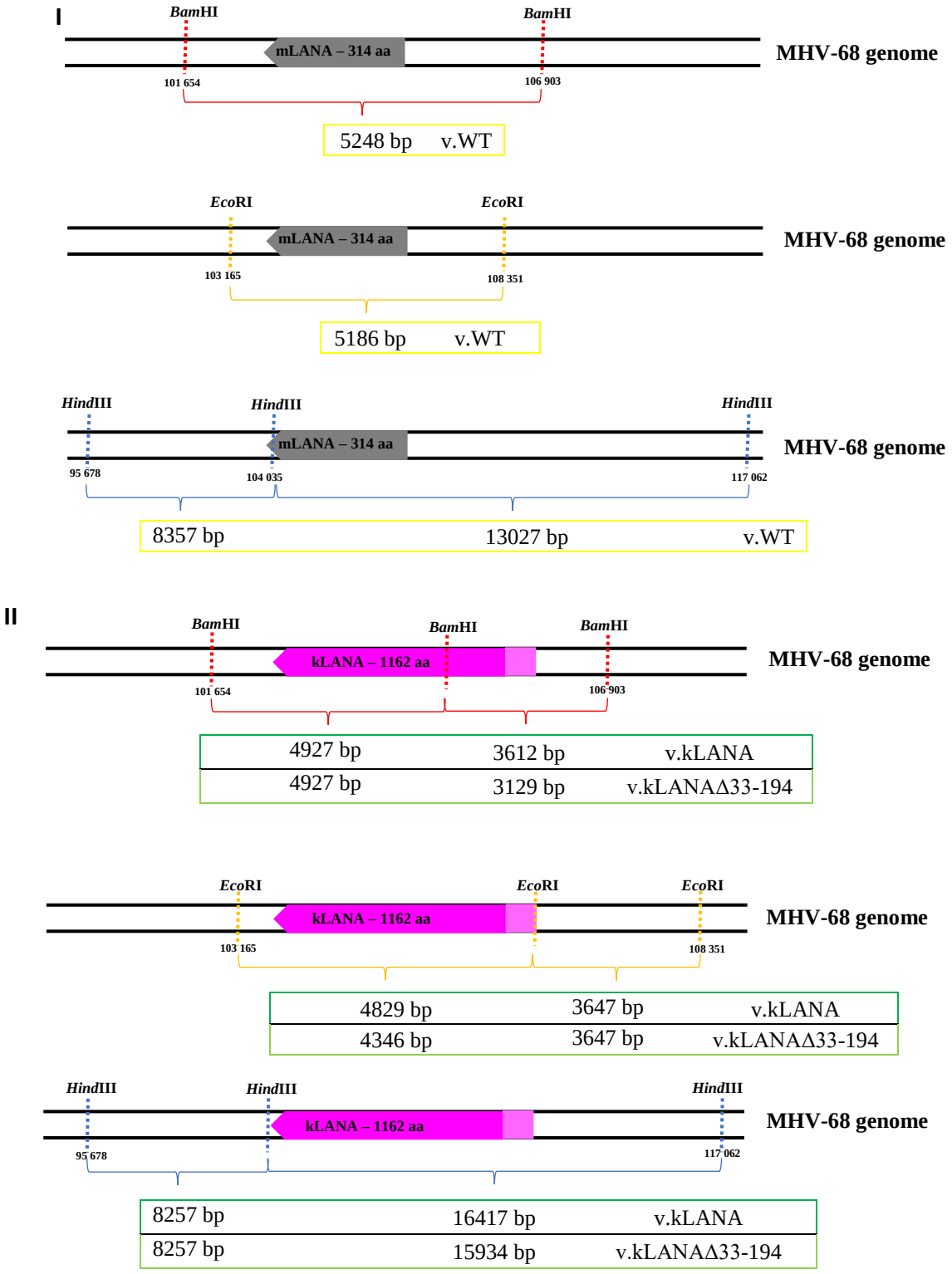


Figure 4.3. Schematic representation of genome depicting each restriction enzyme's sites on (I) v.WT and (II) v.kLANA. The light rose box in v.kLANA corresponds to kLANA 5'UTR.The coordinates indicated in the scheme correspond to MHV-68 WT genome. Figure not to scale.

4.2. *In vitro* assays

4.2.1. Viral titration

Viral stocks produced on BHK-21 fibroblasts were used in all experiments. To determine their concentration, work stock media from each virus was titred (Table 4.1). Even though some lower values, they are all within the range obtained in the laboratory (>10^6 PFU/mL).

Table 4.1. Viral stocks titration. The values are presented in PFU/mL.

Virus	Titer (PFU/mL)
v.WT	6.1×10^6
v.WT.yfp	1.1×10^8
v.kLANA	4.4×10^6
v.kLANA.yfp	3.3×10^6
v.kLANAΔ33-194	3.9×10^6
v.kLANAΔ33-194.yfp	3.8×10^7

4.2.2. Expression of kLANA mutant proteins

To assess if recombinant viruses expressed kLANA properly, BHK-21 cells were infected with each virus at a MOI of 3 PFU/cell, during 6 h. Total cell lysates were prepared and analyzed by immunoblot assay using specific antibodies for kLANA, using as controls mLANA, viral protein M3, eGFP and cellular actin (Figure 4.4.).

Full length and mutant kLANA were detected with a commercial antibody, LN53, which recognizes the repetitive glutamic motifs EQEQE found in the glutamate and glutamine repeat region (EQE). As expected, full length kLANA presented a higher molecular weight (≈250 kDa) than the mutant kLANAΔ33-194 protein (Figure 4.4., top panel). We can also observe several bands that correspond to different isoforms that have been associated to alternative initiation of translation and alternative poly adenylation signal (Toptan *et al*., 2013). mLANA protein, which has a molecular weight of ≈50 kDa, was only detected in WT viruses, as predicted (Figure 4.4., second panel from top). To compare infection levels, an antibody against M3, a chemokine binding protein (≈43 kDa), was used. As demonstrated in figure 4.4. (third panel from top), viruses had similar infection levels even though WT viruses exhibit a lower intensity compared to the other viruses, possible due to slightly lower inoculums. Anti-eGFP was used to detect eGFP protein which is related to YFP protein (≈27 kDa). A signal was observed in all samples infected with yfp viruses (Figure 4.4, last panel). Since there is no signal in samples infected with non.yfp viruses, this also proves that the removal of BAC sequences was successful in these viruses. Actin, the most abundant protein expressed in most eukaryotic cells (≈ 42 kDa) is commonly used as a loading control. All viruses as well as the mock sample presented a blot (Figure 4.4., fourth panel from top) and the level of detection was similar between all.

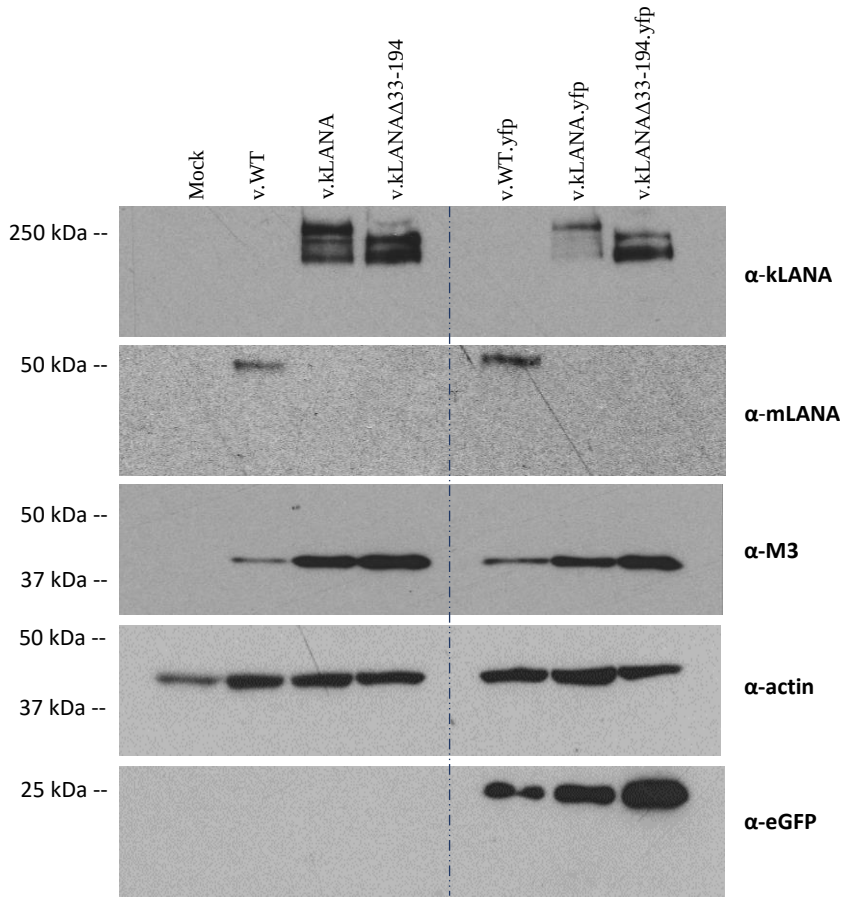


Figure 4.4. Western blot analysis of viral proteins (kLANA, mLANA, M3 and eGFP) and cellular protein (actin) in v.WT, v.kLANA and v.kLANAΔ33-194 in both backgrounds using total cell lysates. 5×10^5 BHK-21 cells were infected with the above-mentioned viruses with a MOI of 3 PFU/cell. A sample from uninfected cells (mock) was used as control. Molecular weight is indicated on the left.

4.2.3. Growth kinetics of recombinant viruses

ORF73 was demonstrated to not be required for an effective lytic replication *in vivo* and *in vitro* (Fowler *et al*., 2003). However, it is important to assess if the deletion introduced on *kLANA* does not show an unexpected result.

To analyze viruses' growth kinetics, a multi-step growth curve was performed. BHK-21 fibroblasts were infected with a low MOI (0.01 PFU/cell). Cells and media were harvested at six specific time-points (0, 24, 48, 72, 96 and 120 h after infection). After freezing and thaw the samples, titrations of all time points were performed by plaque assay and the obtained titers were used to construct a growth curve (Figure 4.5.). v.WT and v.kLANA (from both backgrounds) were used as control viruses.

Mutants displayed a similar growth compared to control viruses. The differences observed on time point 72 h and onwards were not statistically significant (One-way ANOVA).

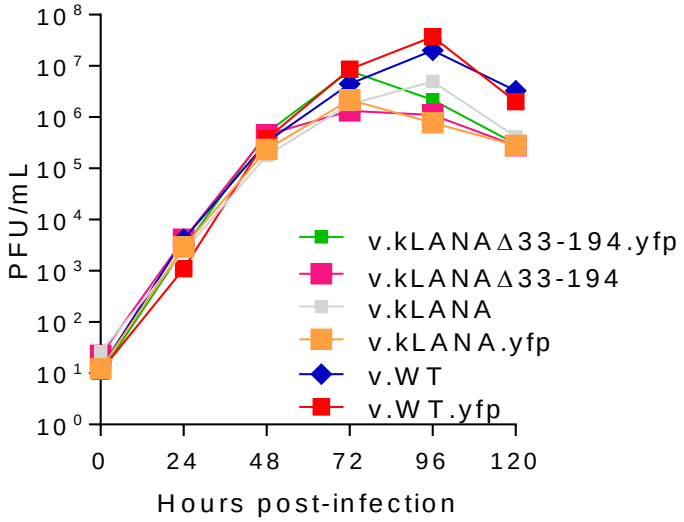


Figure 4.5. Multi-step growth curve of v.WT and chimeric viruses, in both backgrounds. BHK-21 cells were infected with 0.01 PFU/cell. At specific times after infection (0, 24, 48, 72, 96 and 120 h), cells were harvested and freeze-thawed to perform plaque assay to determine viruses' titers.

These results are in accordance with previous results that indicate that the substitution of *mLANA* by *kLANA* does not have a significant impact on lytic viral growth. The results also demonstrate that the deletion in study does not have a significant effect on viral growth kinetics.

4.3. *In vivo* assays

4.3.1. Infectious center assay

Previous studies demonstrated that *mLANA* is essential for the establishment of latent phase in the spleen. *kLANA* partially rescues MHV-68 latency when *mLANA* is replaced by *kLANA* (Moorman *et al*., 2003; Fowler *et al*., 2003; Gupta *et al*., 2017).

The reactivation assay or infectious center assay is based on the reactivation of latent virus (from splenocytes) on permissive BHK-21 cells and subsequent formation of viral plaques. This way, it is possible to quantify the latent viral load in the spleen by analyzing the *ex vivo* reactivation of viruses. C57BL/6J mice were intranasally inoculated with 10^4 PFU and 14 days post-infection they were sacrificed, and their spleens were removed and processed. The reactivation leads to the formation of viral plaques within the monolayer that can be counted and titers per spleen can be calculated (Figure 4.6., closed circles). To analyze the presence of lytic infectious viruses in the spleen, the single cell suspensions from each mouse were freeze-thawed to disrupt cells and were co-cultured with permissive BHK-21 cells. Pre-formed viruses were analyzed by counting the number of viral plaques and results were expressed in PFU per spleen (Figure 4.6., open circles).

As observed, despite the lower latent load, v.kLANA.yfp was able to establish latency in the spleen, which is in accordance to results previously obtained and described (Habison *et al*., 2017). The latency in the v.kLANAΔ33-194.yfp seemed not to be affected compared to v.kLANA.yfp. However, the v.kLANAΔ33-194 mutant appeared to be unable to establish latency, since the results are really close to or below the limit of detection. These results are unexpected since it was demonstrated that the fusion of YFP protein with the MHV-68 genome does not affect the establishment of latency (Collins *et al*., 2009). Therefore, it was expected that both recombinant viruses had similar phenotypes. This may indicate that a random mutation was introduced.

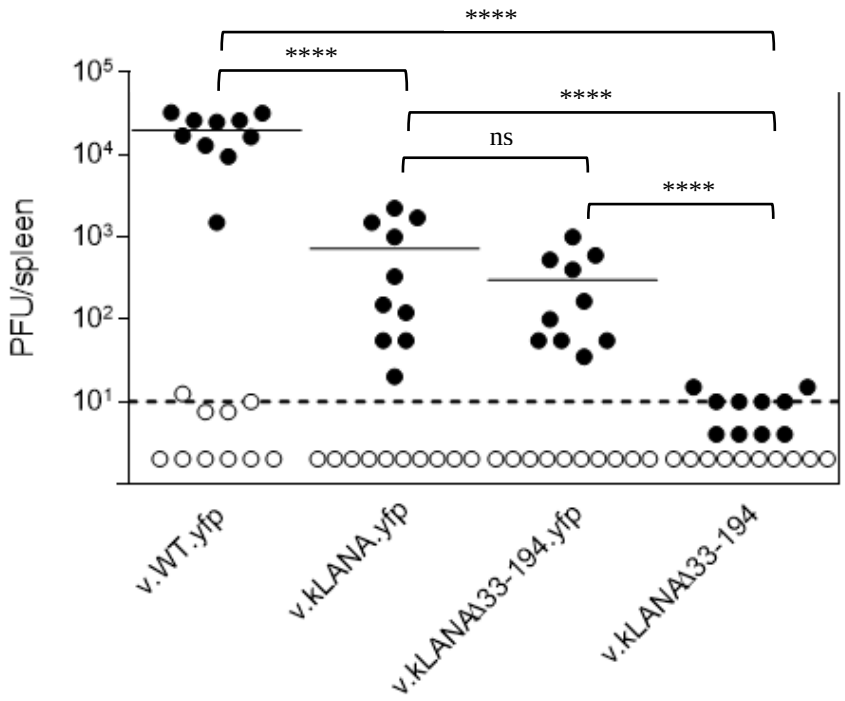


Figure 4.6. Quantification of latent infection (PFU/spleen) in the spleens of mice infected with v.WT.yfp, v.kLANA.yfp, v.kLANAΔ33-194.yfp and v.kLANAΔ33-194, 14 days post-infection, by *ex vivo* reactivation assay. Closed circles correspond to titers of latent viruses. Titers of pre-formed viruses were obtained by freeze-thawed splenocytes suspensions and are represented by open circles. Each circle represents the titer of an individual mouse. The dashed line represents the limit of detection of the assay (10^1 PFU/spleen). The horizontal bars denote mean values. The data presented in this graph is a combination of two independent experiments with five mice per infectious group. **** ($p<0.0001$); ns: non-significant ($p>0.05$).

4.3.2. Quantification of viral DNA positive cells by limiting dilution assay

In addition to the reactivation assay, a limiting dilution assay was also performed to quantify the viral DNA positive cells on total splenocytes. This assay does not depend on the ability of the virus to reactivate *ex vivo*, and thus is an independent measurement of latency. The results were obtained by real time PCR analysis using a probe and specific primers for MHV-68 *M9* gene.

The v.WT.yfp and v.kLANA.yfp viruses exhibited the expected values already described (Habison *et al*., 2017; Pires de Miranda *et al*., 2018). The v.kLANAΔ33-194.yfp showed no difference or a slight difference compared to v.kLANA.yfp. Again, v.kLANAΔ33-194 displayed a significant lower frequency of viral DNA positive cells compared to v.kLANAΔ33-194.yfp (Figure 4.7. A and B) (Table 4.2.). Therefore, the results obtained in the limiting dilution assay are in accordance with the results obtained in the infectious center assay.

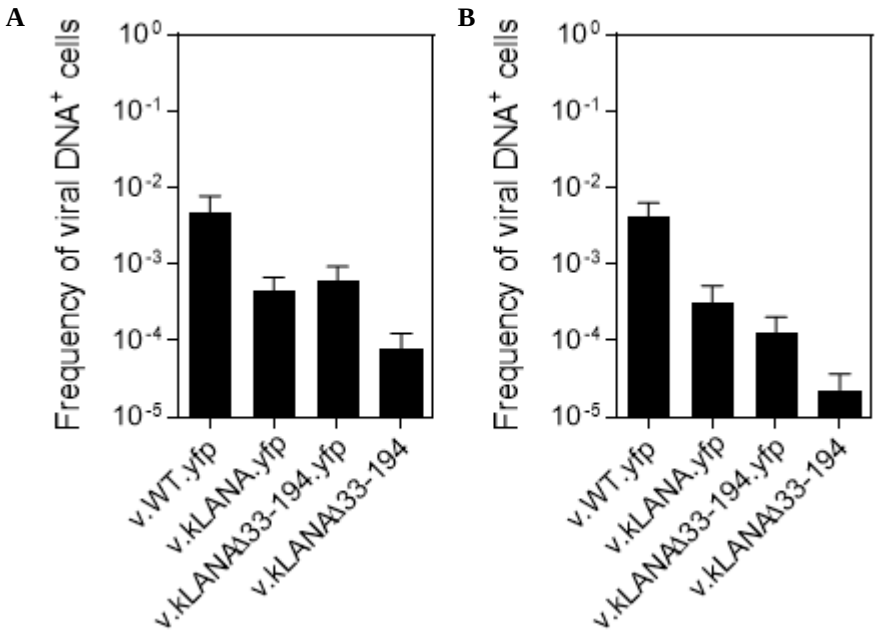


Figure 4.7. Frequency of viral DNA positive cells in total splenocytes, obtained by limiting dilution assay and real-time PCR. The two graphics were obtained from two independent experiments (A and B). Data were obtained from pools of five mice. Bars represent frequency of viral DNA positive cells with 95% confidence intervals.

Table 4.2. Reciprocal frequency of viral DNA positive cells in total splenocytes[2].

Cell population	Experiment	DPI	Virus	Reciprocal frequency of viral DNA positive cells[3]	95% confidence interval
Total splenocytes	A	14	v.WT.yfp	214,48	(128,45-649,50)
			v.kLANA.yfp	2269,33	(1485,51-4804,34)
			v.kLANAΔ33-194.yfp	1682,39	(1071,88-3908,64)
			v.kLANAΔ33-194	12164	(7810-27498)
	B	14	v.WT.yfp	244,52	(157,78-453,07))
			v.kLANA.yfp	3130,73	(1915,04-8572,87)
			v.kLANAΔ33-194.yfp	7736,78	(4893,6-18464,9)
			v.kLANAΔ33-194	44001,88	(27356,2-112387,4)

[2] Data were obtained from pools of five mice.
[3] Frequencies were calculated by limiting dilution analysis with 95% confidence intervals.

4.3.3. Flow cytometry analysis

For the establishment of latency, MHV-68 has GC B cells as the main target. By performing flow cytometry, it is possible to quantify the latent infection in the B cell population of the spleen. Splenocyte suspensions from v.WT.yfp, v.kLANA.yfp and v.kLANAΔ33-194.yfp were incubated with the suitable antibodies to detect germinal center B cells. The YFP protein expressed by v.WT.yfp, v.kLANA.yfp and v.kLANAΔ33-194.yfp allows the determination of infected cells. Therefore, the frequency of GC B cells that are latently infected by the respective viruses was determined by flow cytometry analysis.

The percentage of B cells present in germinal centers ($CD19^+GL7^+$) varied between 3 and 6%, 1 and 7%, 3 and 7% in v.WT.yfp, v.kLANA.yfp and v.kLANAΔ33-194.yfp, respectively. Considering average percentages, no statistically significant differences were observed between the three infectious groups (Figure 4.8.).

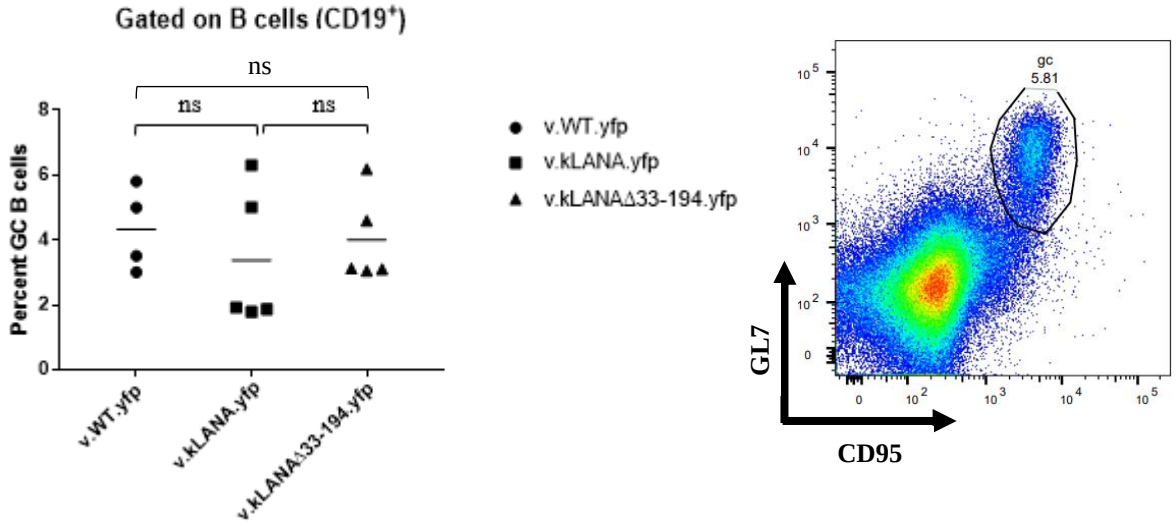


Figure 4.8. Flow cytometry analysis. Percentage quantification of GC B cells (CD19+). Each dot represents an individual mouse and the horizontal bar denotes mean values. Representative FACS dot plot gated on B cells is shown on the right. ns: non-significant ($p>0.05$).

Given the expression of YFP protein by all infectious groups, it is possible to quantify the frequency of infected GC B cells. Results showed that v.WT.yfp had a frequency of 11.46%, v.kLANA.yfp presented an average percentage of 1.97% and v.kLANAΔ33-194.yfp had a frequency of 0.71% (Figure 4.9.). The difference observed between v.kLANA.yfp and the recombinant virus might be due to the decline of YFP signal after the onset of latency. This leads to a reduction of the number of infected B cells with YFP signal (Collins and Speck, 2012). Nonetheless, this experiment needs to be repeated to increase the significance of the results.

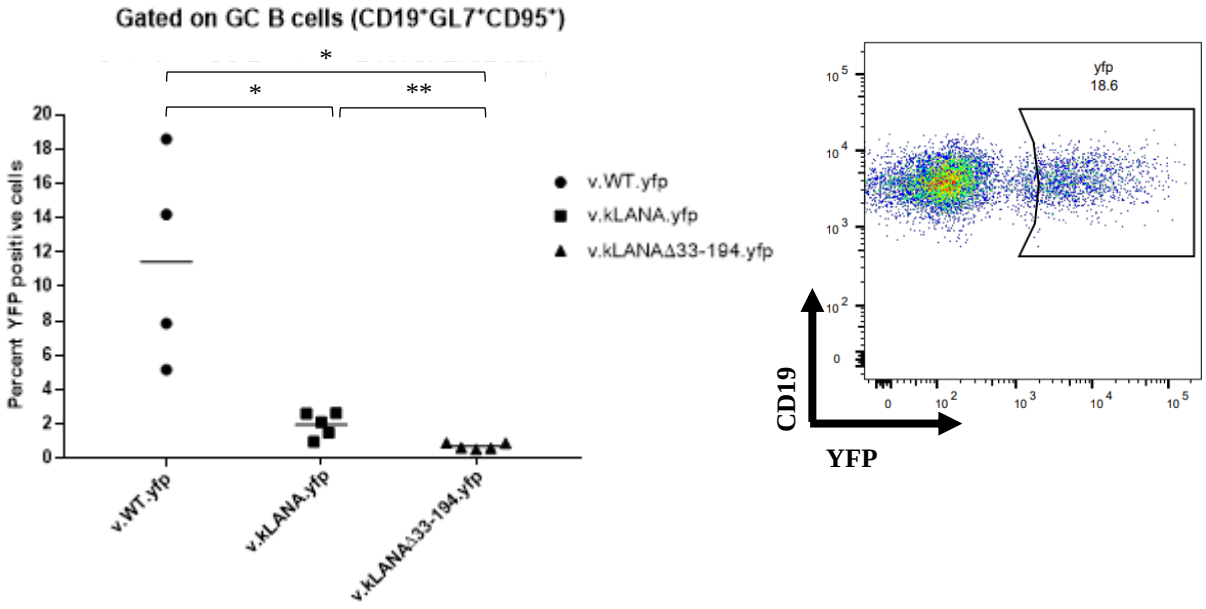


Figure 4.9. Flow cytometry analysis. Percentage quantification of YFP positive (infected) GC B cells (CD19+GL7+CD95+). Each dot represents an individual mouse and the horizontal bar denotes mean values. Representative FACS dot plot gated on GC B cells is shown on the right. * ($p<0.05$); ** ($p<0.01$).

Lastly, the frequency of infected B cells with GC phenotype was quantified. v.WT.yfp, v.kLANA.yfp and v.kLANAΔ33-194.yfp presented 80.05%, 42.4% and 29.99%, respectively, of infected B cells with GC phenotype (figure 4.10.). No statistically significant differences were observed between v.kLANA.yfp and v.kLANAΔ33-194.yfp.

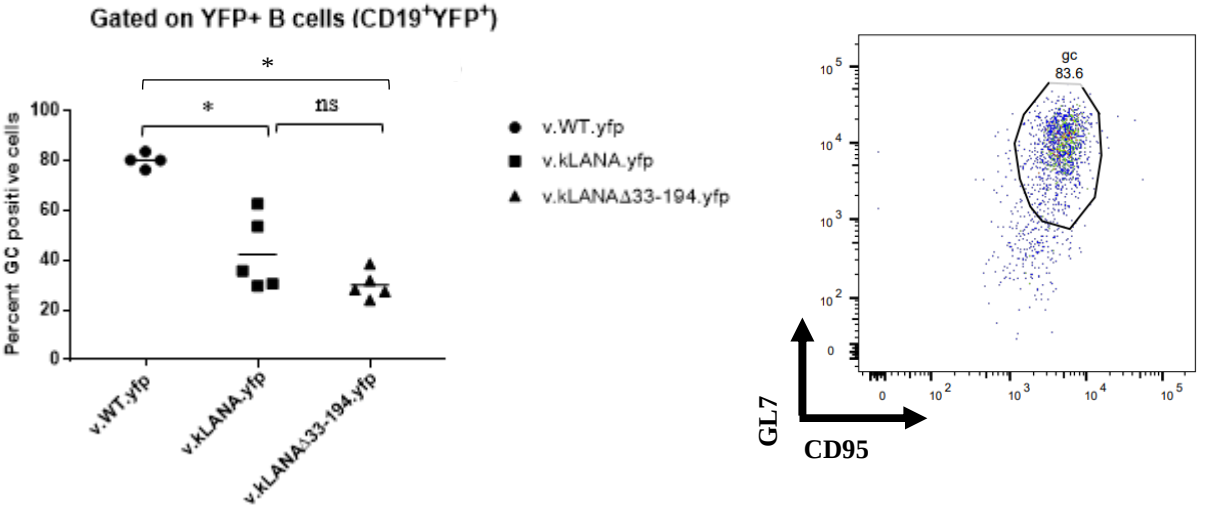


Figure 4.10. Flow cytometry analysis. Percentage quantification of YFP positive (infected) B cells with GC phenotype (CD19+YFP+). Each dot represents an individual mouse and the horizontal bar denotes mean values. Representative FACS dot plot gated on YFP positive B cells is shown on the right. * ($p<0.05$); ns: non-significant ($p>0.05$).

5. Discussion and future perspectives

Kaposi sarcoma-associated herpesvirus is a gammaherpesvirus related to several diseases that is able to establish a latent phase within host's B cells. This phase is characterized by circularization of viral genome and a highly restricted gene expression. Among the proteins expressed, the LANA protein was shown to be essential for the replication and segregation of the episome, functions that are crucial for the maintenance of the virus. A better understating of the functionality of this protein is essential to the development of strategies to control KSHV latent infection.

In vitro studies demonstrated that the region of kLANA encompassing amino acids 33 to 194 is essential for the segregation of viral episome.

In this study, a recombinant virus was constructed with a deletion between amino acids 33 to 194 within *kLANA* to assess the impact of the mutation on the establishment of latency using an *in vivo* model of infection. The recombinant virus contains MHV-68 genome but *mLANA* was replaced by *kLANA* containing the deletion. The virus was also generated in a YFP background that allows the direct identification of infected cells.

In vitro assays were performed to verify protein expression and viral growth kinetics. As expected, by immunoblotting assay it was possible to observe that the recombinant viruses expressed the LANA mutant protein with a lower molecular weight. Therefore, the deletion does not affect the expression of LANA protein. Results from multi-step growth curve showed that the recombinant viruses display a normal growth kinetics indicating that neither the replacement of *mLANA* by *kLANA* nor the introduction of the deletion significantly affect *in vitro* lytic replication.

Results from *in vivo* experiments showed a difference in latency levels between v.WT.yfp and v.kLANA.yfp, as already described by Habison *et al* (2017). The v.kLANAΔ33-194.yfp presented similar levels of latency compared to v.kLANA.yfp in the reactivation assay which were corroborated by limiting dilution assay. However, comparing to v.kLANAΔ33-194.yfp, v.kLANAΔ33-194 could not establish latency and had a lower frequency of viral DNA positive cells. This significant difference between the two recombinant viruses was not expected since the coupling of YFP protein into the MHV-68 genome has been demonstrated to not interfere with latency levels (Collins *et al.*, 2009). It suggests that a random mutation was introduced possibly during the mutagenesis procedure, but it could also have occurred during the removal of BAC sequences or even during viral propagation.

Given this problem, the importance of the region is still unclear. Therefore, further studies must be done to overcome the problem and taking out a conclusion. These may include analysis of more clones or the construction of a revertant virus in which *kLANAΔ33-194* from the recombinant virus is replaced by *mLANA*. Other experiments such as analysis of lytic infection in lungs, analysis of latent infection in the spleen at later stages or examination of other secondary lymphoid organs may also be convenient to better characterize the phenotype of this mutant. Besides that, protein interaction analysis might be useful to possibly map certain proteins to that region.

Future experiments to construct recombinant viruses may consider the use of modern gene editing strategies. To save time and avoid the removal of unwanted sequences (such as BAC sequences), CRISPR-Cas9 (clustered regularly interspaced short palindromic repeats associated cas9 protein) might be a more accurate and efficient alternative for the introduction of the desired mutations (Suenaga *et al*., 2014). The CRISPR-Cas9 type II system is the most studied mechanism and has been adapted from the naturally occurring system in bacteria to protect them against phages. It comprises three steps: adaptation, expression and interference. First, when an infection occurs, segments of foreign DNA are

incorporated into the CRISPR locus within bacterial genome. When a new infection takes place, the CRISPR locus is transcribed (crRNA) and this RNA molecule together with tracrRNA (non-coding trans-activating crRNA) and Cas9 protein form a complex that binds the target (which is adjacent to the protospacer adjacent motif (PAM)) and cleaves the foreign DNA (Rath *et al*., 2015). The simplicity of this system makes it useful for genome editing. Engineered CRISPR-Cas9 system is compound by two components: a single-guided RNA (sgRNA) which is a RNA molecule that includes crRNA and tracrRNA, and a Cas9 endonuclease. The target can be any 20 nucleotides sequence as long as it is unique within the genome and it has a PAM next to the target. The sgRNA and Cas9 protein form a ribonucleoprotein complex that will bind the target through a zipper-like annealing mechanism. Cas9 HNH domain cleaves one strand and RuvC domain will cleave the other strand, forming a double strand break within the target. Once the DNA is cut, the repair mechanisms are activated and include non-homologous end joining or homologous direct repair (when a homologous repair template is present). Several strategies can be used to optimize the technique according to the objective of the experiment.

Recent studies using the CRISPR-Cas9 technique as a therapeutic approach allowed the excision of few genes including EBNA1 of EBV (protein functionally equivalent to KSHV LANA). For this, two sgRNAs were constructed to both ends of EBNA1 coding region and plasmids containing those sgRNAs and Cas9 protein coupled with EGFP marker were transfected into cells. Cells which were successfully transfected with the system were identified and sorted by flow cytometry. The deletion was further confirmed by PCR and sequencing (Wang and Quake, 2014). The use of CRISPR-Cas9 in cells infected with latent EBV may shed a light that the system may also work with KSHV and MHV-68. Extrapolating to our study, a cell line stably harboring MHV-68 genome would be necessary to be transfected with two sgRNA and *cas9* gene. These are necessary to excise *mLANA*, and a donor template with *kLANA* encompassing the desired mutation would also be required. One method to validate the process consist in the analysis of the restriction profile or the analysis of the PCR product using primers flanking the modified region. The next step would involve the generation of clonal cell lines by serial dilutions to isolate single cells. After expansion, cell lines must be tested through PCR amplification and/or sequencing and protein expression may also be analyzed.

Besides that, the lower titers observed by the *ex vivo* reactivation assay can be improved using another mouse strain such as BALB/c. Preliminary results from a recent work developed in our laboratory demonstrated that infection of BALB/c mice with v.WT.yfp and v.kLANA.yfp led to higher levels of latent infection in the spleen, 14 days post-infection, compared to infection of C57BL/6J mice with the same viruses. Nevertheless, the difference in latency levels between v.WT.yfp and v.kLANA.yfp is also observed in BALB/c mice. Given these results, experiments with BALB/c mice may afford more expressive results. However, more experiments should be done to confirm and explore these results (unpublished data).

In conclusion, additional work must be done to better understand LANA protein functionality that will subsequently lead to new insights into viral latency. This may provide new targets to control latency and to treat KSHV infection.

6. References

Adler, H., Messerle, M., Wagner, M., and Koszinowski, U. H. (2000). Cloning and Mutagenesis of the Murine Gammaherpesvirus 68 Genome as an Infectious Bacterial Artificial Chromosome. *J Virol*, 74(15), pp: 6964–6974;

Adler, H., Messerle, M. and Koszinowski, U. H. (2001) Virus reconstituted from infectious bacterial chromosome (BAC)-cloned murine gammaherpesvirus 68 acquires wild-type properties *in vivo* only after excision of BAC vector sequences. *J Virol*, 75(12), pp: 5692-5696;

Avey, D., Brewers, B. and Zhu, F. (2015). Recent advances in the study of Kaposi's sarcoma-associated herpesvirus replication and pathogenesis. *Virol Sin*, 30(2), pp: 130-145;

Ballestas, M. E., Chatis, P. A. and Kaye, K. M. (1999). Efficient persistence of extrachromosomal KSHV DNA mediated by latency-associated nuclear antigen. *Science*. 284(5414), pp: 641-644;

Ballestas, M. E. and Kaye, K. M. (2001). Kaposi's Sarcoma-Associated Herpesvirus Latency-Associated Nuclear Antigen 1 Mediates Episome Persistence through cis-Acting Terminal Repeat (TR) Sequence and Specifically Binds TR DNA. *J Virol*, 75(7), pp: 3250–3258;

Barbera, A. J., Chodaparambil, J. V., Kelley-Clarke, B., Joukov, V., Walter, J. C., Luger, K. and Kaye, K. M. (2006). The nucleosomal surface as a docking station for Kaposi's sarcoma herpesvirus LANA. *Science*, 311(5762), pp: 856-861;

Barton, E., Mandal, P. and Speck, S. H. (2011). Pathogenesis and host control of gammherpesviruses: lessons from the mouse. *Annu Rev Immunol*, 29, pp: 351-397;

Blasdell, K., McCracken, C., Morris, A., Nash, A., Begon, M., Bennett, M. and Stewart, J. (2003) The wood mouse is a natural host for Murid herpesvirus 4. *J Gen Virol*, 84(1), pp:111-113;

Blaskovic, D., Stancekova, M., Svobodova, J. and Mistrikova, J. (1980). Isolation of five strains of herpesviruses from two species of free-living small rodents. *Acta Virol*, 24(6), 468;

Bonnefoix, T., Bonnefoix, P., Callanan, M., Verdiel, P. and Sotto, J. J. (2001). Graphical representation of a generalized linear model-based statistical test estimating the fit of the single-hit Poisson model to limiting dilution assays. *J Immunol*, 167(10), pp: 5725-5730;

Burrell, C. J., Howard, C. R. and Murphy, F. A. (2017). Herpesvirus. In: Burrell, C. J., Howard, C. R. and Murphy, F. A. Fenner and White's Medical Virology. (pp: 237-261). Academic Press;

Cai, Q., Verma, S. C., Lu, J. and Robertson, E. S. (2010). Molecular Biology of Kaposi's Sarcoma Herpesvirus and Related Oncogenesis. *Adv Virus Res*, 78, pp: 87–142;

Cao, W., Vyboh, K., Routy, B., Chababi-Atallah, M., Lemire, B. and Routy, J. P. (2015). Imatinib for highly chemoresistant Kaposi sarcoma in a patient with long-term HIV control: a case report and literature review. *Curr Oncol*, 22(5), pp: 395–399;

Chakraborty, S., Veettil, M. V. and Chandran, B. (2012). Kaposi's Sarcoma Associated Herpesvirus Entry into Target Cells. *Front Microbiol*, 3(6);

Chang, Y., Cesarman, E., Pessin, M. S., Lee, F., Culpepper, J., Knowles, D. M. and Moore, P. S. (1994). Identification of herpesvirus-like DNA sequences in AIDS-associated Kaposi's sarcoma. *Science*, 266(5192), pp: 1865-1869;

Collins, C. M., Boss, J. M. and Speck, S. H. (2009). Identification of Infected B-Cell Populations by Using a Recombinant Murine Gammaherpesvirus 68 Expressing a Fluorescent Protein. *J Virol*, 83(13), pp: 6484–6493;

Collins, C. M. and Speck, S. H. (2012). Tracking Murine Gammaherpesvirus 68 Infection of Germinal Center B Cells *In Vivo*. *PLoS One*, 7(3), e:33230;

Correia, B., Cerqueira, S. A., Beauchemin, C., Pires de Miranda, M., Li, S., Ponnusamy, R., Rodrigues, L., Schneider, T. R., Carrondo, M. A., Kaye, K. M., Simas, J.P. and McVey, C. E. (2013) Crystal Structure of the Gamma-2 Herpesvirus LANA DNA Binding Domain Identifies Charged Surface Residues Which Impact Viral Latency. *PLoS Pathog*, 9(10), e:1003673;

Coscoy, L. (2007). Immune evasion by Kaposi's sarcoma-associated herpesvirus. *Nat Rev Immunol*, 7(5), pp: 391-401;

Davison, A. J. (2002). Evolution of the herpesviruses. *Vet microbiol*, 86(1-2), pp: 69–88;

Davison, A. J., Eberle, R., Ehlers, B., Hayward, G. S., McGeoch, D. J., Minson, A. C., Pellett, P. E., Roizman, B., Studdert, M. J. and Thiry, E. (2009). The Order *Herpesvirales*. *Arch Virol*, 154(1), pp: 171-177;

De León Vázquez, E., Carey, V. J. and Kaye, K. M. (2013). Identification of Kaposi's Sarcoma-Associated Herpesvirus LANA Regions Important for Episome Segregation, Replication, and Persistence. *J Virol*, 87(22), pp: 12270-12283;

De León Vázquez, E., Juillard, F., Rosner, B. and Kaye, K. M. (2014). A short sequence immediately upstream of the internal repeat elements is critical for KSHV LANA mediated DNA replication and impacts episome persistence. *Virology*, 448, pp: 344-355;

Efstathiou, S., Ho, Y. M. and Minson, A. C. (1990). Cloning and molecular characterization of the murine herpesvirus 68 genome. *J Gen Virol*, 71(6), pp:1355-1364;

Flaño, E., Husain, S. M., Sample, J. T., Woodland, D. L. and Blackman, M. A. (2000). Latent murine γ-Herpesvirus Infection Is Established in Activated B Cells, Dendritic Cells, and Macrophages. *J Immunol*, 165(2), pp: 1074-1081;

Fowler, P., Marques, S., Simas, J. P. and Efstathiou, S. (2003). ORF73 of murine herpesvirus-68 is critical for the establishment and maintenance of latency. *J Gen Virol*, 84(12), pp: 3405-3412;

Frederico, B., Chao, B., May, J. S., Belz, G. T. and Stevenson, P. G. (2014). A murid gamma-herpesviruses exploits normal splenic immune communication routes for systemic spread. *Cell Host microbe*, 15(4), pp: 457-470;

Garber, A. C., Hu, J. and Renne, R. (2002). Latency-associated Nuclear Antigen (LANA) Cooperatively Binds to Two Sites within the Terminal Repeat, and Both Sites Contribute to the Ability of LANA to Suppress Transcription and to Facilitate DNA Replication. *J Biol Chem*, 277(30), pp: 27401-27411;

Giffin, L. and Damania, B. (2014). KSHV: Pathways to Tumorigenesis and Persistent Infection. *Adv Virus Res*, 88, pp: 111-159;

Godinho-Silva, C., Marques, S., Fontinha, D., Veiga-Fernandes, H, Stevenson, P. G. and Simas, J. P. (2014). Defining Immune engagement thresholds for in vivo control of virus-driven lymphoproliferation. *Plos Pathog*, 10(6), e:1004220;

Grinde, B. (2013). Herpesviruses: latency and reactivation – viral strategies and host response. *J Oral Microbiol*, 5;

Gupta, A., Oldenburg, D. G., Salinas, E., White, D. W. and Forrest, J. C. (2017). Murine Gammaherpesvirus 68 Expressing Kaposi Sarcoma-Associated Herpesvirus Latency-Associated Nuclear Antigen (LANA) Reveals both Functional Conservation and Divergence in LANA Homologs. *J Virol*, 91(19), e00992–17;

Habison, A. C., Beauchemin, C., Simas, J. P., Usherwood, E. J. and Kaye, K. M. (2012). Murine Gammaherpesvirus 68 LANA Acts on Terminal Repeat DNA To Mediate Episome Persistence. *J Virol*, 86(21), pp: 11863–11876;

Habison, A. C., de Miranda, M.P., Beauchemin, C., Tan, M., Cerqueira, S. A., Correia, B. and Kaye, K. M. (2017). Cross-species conservation of episome maintenance provides a basis for in vivo investigation of Kaposi's sarcoma herpesvirus LANA. *PLoS Pathog*, 13(9), e:1006555;

Hellert, J., Weidner-Glunde, M., Krausze, J., Lünsdorf, H., Ritter, C., Schulz, T. F. and Lührs, T. (2015) The 3D structure of Kaposi sarcoma herpesvirus LANA C-terminal domain bound to DNA. *Proc Natl Acad Sci U S A*, 112(21), pp: 6694-6699;

Hu, J. and Renne, R. (2005). Characterization of the Minimal Replicator of Kaposi's Sarcoma-Associated Herpesvirus Latent Origin. *J Virol*, 79(4), pp: 2637–2642;

Jensen, K.K., Chen, S.C., Hipkin, R.W., Wiekowski, M.T., Schwarz, M.A., Chou, C.C., Simas, J.P., Alcami, A. and Lira, S.A. (2003). Disruption of CCL21-induced chemotaxis in vitro and in vivo by M3, a chemokine-binding protein encoded by murine gammaherpesvirus 68. J Virol, 77(1), pp:624–630;

Juillard, F., Tan, M., Li, S., and Kaye, K. M. (2016). Kaposi's Sarcoma Herpesvirus Genome Persistence. *Front Microbiol*, 7, 1149;

Kwun, H. J., Silva, S. R., Shah, I. M, Blake, N., Moore, P. S. and Chang, Y. (2007). Kaposi's sarcoma-associated herpesviruses latency-associated nuclear antigen 1 mimics Epstein-barr virus EBNA1 immune evasion through central repeat domain effects on protein processing. *J Virol*, 81(15), pp: 8225-8235;

Li, S., Tan, M., Juillard, F., Ponnusamy, R., Correia, B., Simas, J. P., Carrondo, M. A., McVey, C. E. and Kaye, K. M. (2015). The Kaposi Sarcoma Herpesvirus Latency-associated Nuclear Antigen DNA Binding Domain Dorsal Positive Electrostatic Patch Facilitates DNA Replication and Episome Persistence. *J Biol Chem*, 290(47), pp: 28084–28096;

Lin, C. L., Li, H., Wang, Y., Zhu, F. X., Kudchodkar, S. and Yuan, Y. (2003). Kaposi's Sarcoma-Associated Herpesvirus Lytic Origin (*ori-Lyt*)-Dependent DNA Replication: Identification of the *ori-Lyt* and Association of K8 bZip Protein with the Origin. *J Virol*, 77(10), pp:5578–5588;

Liu, F. and Zhou, Z. H. (2007). Comparative virion structures of human herpesviruses. In: Arvin A, Campadelli-Fiume G, Mocarski E, *et al*., editors. Human Herpesviruses: Biology, Therapy, and Immunoprophylaxis. Cambridge: Cambridge University Press;

Louten, J. (2016) Herpesviruses. In: Louten, J. Essential Human Virology. (pp: 235-256). Academic Press;

Mariggiò, G., Koch, S. and Schulz, T. F. (2017). Kaposi sarcoma herpesvirus pathogenesis. *Philos Trans R Soc Lond B Biol Sci*, 372(1732);

Marques, S., Efstathiou, S., Smith, K. G., Haury, M. and Simas, J. P. (2003). Selective gene expression of latent murine gammaherpesvirus 68 in B lymphocytes. *J Virol*, 77(13), pp: 7308-7318;

Mesri, E. A., Cesarman, E. and Boshoff, C. (2010). Kaposi's sarcoma herpesvirus/ Human herpesvirus-8 (KSHV/HHV8), and the oncogenesis of Kaposi's sarcoma. *Nat Rev Cancer*, 10(10), pp: 707–719;

Moore, P. S. and Chang, Y. (2003). Kaposi's sarcoma–associated herpesvirus immunoevasion and tumorigenesis: two sides of the same coin? *Annu Rev Microbiol*, 57, pp: 609–639;

Moorman, N. J., Willer, D. O. and Speck, S. H. (2003). The Gammaherpesvirus 68 Latency-Associated Nuclear Antigen Homolog Is Critical for the Establishment of Splenic Latency. *J Virol*, 77(19), pp: 10295-10303;

Mui, U. N., Haley, C. T. and Tyring, S. K. (2017). Viral oncology: molecular biology and pathogenesis. *J Clin Med*, 6(12);

Nash, A. A., Dutia, B. M., Stewart, J. P., and Davison, A. J. (2001). Natural history of murine gamma-herpesvirus infection. *Philos Trans R Soc Lon B Biol Sci*, 356(1408), pp: 569–579;

O'Connor, M., Peifer, M. and Bender, W. (1989). Construction of large DNA segments in Escherichia coli. *Science*, 244(4910), pp: 1307-1312;

Ohsaki, E. and Ueda, K. (2012). Kaposi's sarcoma-associated herpesvirus genome replication, partitioning, and maintenance in latency. *Front microbiol*, 3, 7;

Piolot, T., Tramier, M., Coppey, M., Nicolas, J. C. and Marechal, V. (2001). Close but distinct regions of human herpesvirus 8 latency-associated nuclear antigen 1 are responsible for nuclear targeting and binding to human mitotic chromosomes. *J Virol*, 75(8), pp: 83948-83959;

Pires de Miranda, M., Lopes, F. B., McVey, C. E., Bustelo, X. R. and Simas, J. P. (2013). Role of Src Homology Domain Binding in Signaling Complexes Assembled by the Murid γ-Herpesvirus M2 Protein. *The Journal of Biological Chemistry*, 288(6), pp: 3858–3870;

Pires de Miranda, M., Quendera, A. P., McVey, C. E., Kaye, K. M. and Simas, J. P. (2018). In vivo persistence of chimeric vírus after substitution of the KSHV LANA DNA binding domain with that of MuHV-4. *J Virol*, JVI.01251-18;

Ponnusamy, R., Petoukhov, M. V., Correia, B., Custodio, T. F., Juillard, F., Tan, M., Pires de Miranda, M., Carrondo, M. A., Simas, J. P., Kaye, K. M., Svergun, D. I. and McVey, C. E. (2015). KSHV but not MHV-68 LANA induces a strong bend upon binding to terminal repeat viral DNA. *Nucleic Acids Res*, 43(20), pp: 10039–10054;

Purushothaman, P., Uppal, T. and Verma, S. C. (2015). Molecular Biology of KSHV Lytic Reactivation. *Viruses*, 7(1), pp: 116-153;

Purushothaman, P., Dabral, P., Gupta, N., Sarkar, R. and Verma, S. C. (2016). KSHV Genome Replication and Maintenance. *Front Microbiol*, 7, 54;

Rath, D., Amlinger, L., Rath, A. and Lundgren, M. (2015). The CRIPSR-Cas immune system: Biology, mechanisms and application. *Biochimie*, 117, pp: 119-128;

Rohner, E., Wyss, N., Trelle, S., Mbulaiteye, S. M., Egger, M., Novak, U., Zwahlen, M. and Bohlius, J. (2014). HHV-8 seroprevalence: a global view. *Syst Rev*, 3, 11;

Russo, J. J., Bohenzky, R. A., Chien, M. C., Chen, J., Yan, M., Maddalena, D., Parry, J. P., Peruzzi, D., Edelman, I. S., Chang, Y. and Moore, P. S. (1996). Nucleotide sequence of the Kaposi sarcoma-associated herpesvirus (HHV8). *Proc Natl Acad Sci U S A*, 93(25), pp: 14862-14867;

Simas, J. P. and Efstathiou, S. (1998). Murine gammaherpesvirus 68: a model for the study of gammaherpesvirus pathogenesis. *Trends Microbiol*, 6(7), pp: 276-282;

Spear, P. G. and Longnecker, R. (2003). Herpesvirus Entry: an update. *J Virol*, 77(19), pp: 10179–10185;

Speck, S. H. and Ganem, D. (2010). Viral latency and its regulation: lessons from the gammaherpesviruses. *Cell Host Microbe*, 8(1), pp: 100–115;

Suenaga, T., Kohyama, M., Hirayasu, K. and Arase, H. (2014). Engineering large viral DNA genomes using the CRISPR-Cas9 system. *Microbiol Immunol*, 58(9), pp: 513-522;

Sun, Q., Tsurimoto, T., Juillard, F., Li, L., Li, S., De León Vázquez, E., Chen, S. and Kaye, K. (2014). Kaposi's sarcoma-associated herpesvirus LANA recruits the DNA polymerase clamp loader to mediate efficient replication and virus persistence. *Proc Natl Acad Sci USA*, 111(32), pp: 11816-11821;

Sunil-Chandra, N. P., Efstathiou, S., Arno, J. and Nash, A. A. (1992). Virological and pathological features of mice infected with murine gammaherpesvirus 68. *J Gen Virol*, 73(9), pp: 2347-2356;

Thakker, S. and Verma, S. C. (2016). Co-infections and Pathogenesis of KSHV-Associated Malignancies. *Front Microbiol*, 7, 151;

Toptan, T., Fonseca, L., Kwun, H. J., Chang, Y. and Moore, P. S. (2013). Complex alternative cytoplasmic protein isoforms of the Kaposi's sarcoma-associated herpesvirus latency-associated nuclear antigen 1 generated through noncanonical translation initiation. *J Virol*, 87(5), pp: 2744–2755;

Uppal, T., Banerjee, S., Sun, Z., Verma, S. C. and Robertson, E. S. (2014). KSHV LANA – The master regulator of KSHV latency. *Viruses*, 6(12), pp: 4961-4998;

Uppal, T., Jha, H. C., Verma, S. C. and Roberston, E. S. (2015). Chromatinization of the KSHV genome during the KSHV life cycle. *Cancers*, 7(1), pp: 112-142;

Virgin, H. W., Latreille, P., Wamsley, P., Hallsworth, K., Weck, K. E., Dal Canto, A. J. and Speck, S. H. (1997). Complete sequence and genomic analysis of murine gammaherpesvirus 68. *J Virol*, 71(8), pp: 5894-5904;

Wang, J. and Quake, S. R. (2014). RNA-guided endonuclease provides a therapeutic strategy to cure latent herpesviridae infection. *Proc Natl Acad Sci U S A*, 111(36), pp: 13157–13162;

Weidner-Glunde, M., Mariggiò, G. and Schulz, T. F. (2017). Kaposi's Sarcom-Associated herpesvirus latency-associated nuclear antigen: replicating and shielding viral DNA during viral persistence. *J Virol*, 91(14);

Wen, K. W. and Damania, B. (2010). Kaposi sarcoma-associated herpesvirus (KSHV): Molecular biology and oncogenesis. *Cancer Lett*, 289(2), pp.140–150;

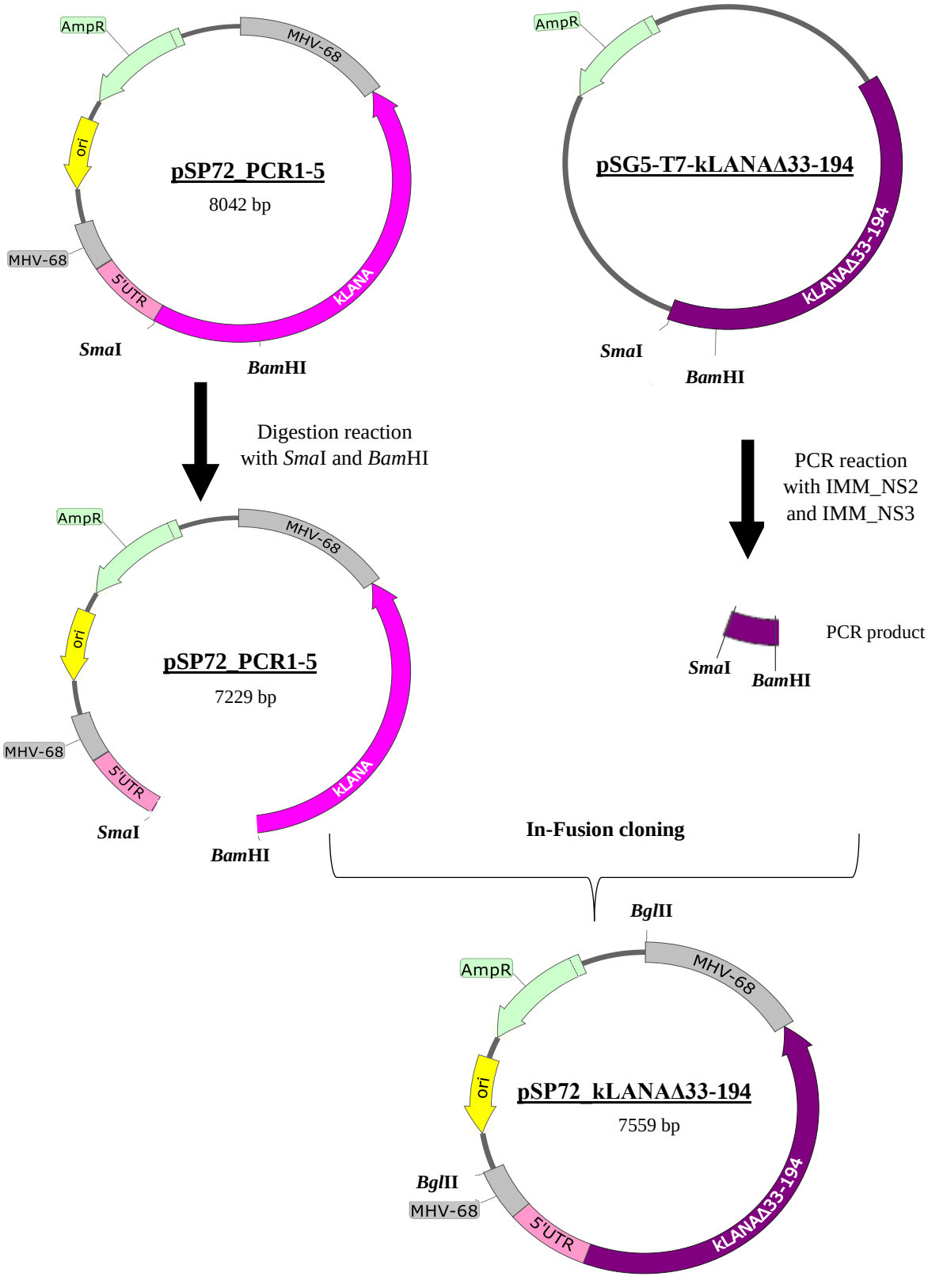
Whitley, R. J. (1996). Herpesviruses. In: Baron S, editor. Medical Microbiology. Galveston (TX): University of Texas Medical Branch at Galveston;

Xiao, B., Verma, S. C., Cai, Q., Kaul, R., Lu, J., Saha, A. and Robertson, E. S. (2010). Bub1 and CENP-F Can Contribute to Kaposi's Sarcoma-Associated Herpesvirus Genome Persistence by Targeting LANA to Kinetochores. *J Virol*, 84(19), pp: 9718–9732;

Xiao, Y., Chen, J., Liao, Q., Wu, Y., Peng, C. and Chen, X. (2013). Lytic infection of Kaposi's sarcoma-associated herpesvirus induces DNA double-strand breaks and impairs non-homologous end joining. *Journal of General Virology*, 94, pp: 1870-1875;

Ye, F.C., Zhou, F.C., Min Yoo, S., Xie, J.P., Browning, P. J. and Gao, S.J. (2004). Disruption of Kaposi's Sarcoma-Associated Herpesvirus Latent Nuclear Antigen Leads to Abortive Episome Persistence. *J Virol*, 78(20), pp: 11121–11129.

Supplementary Data



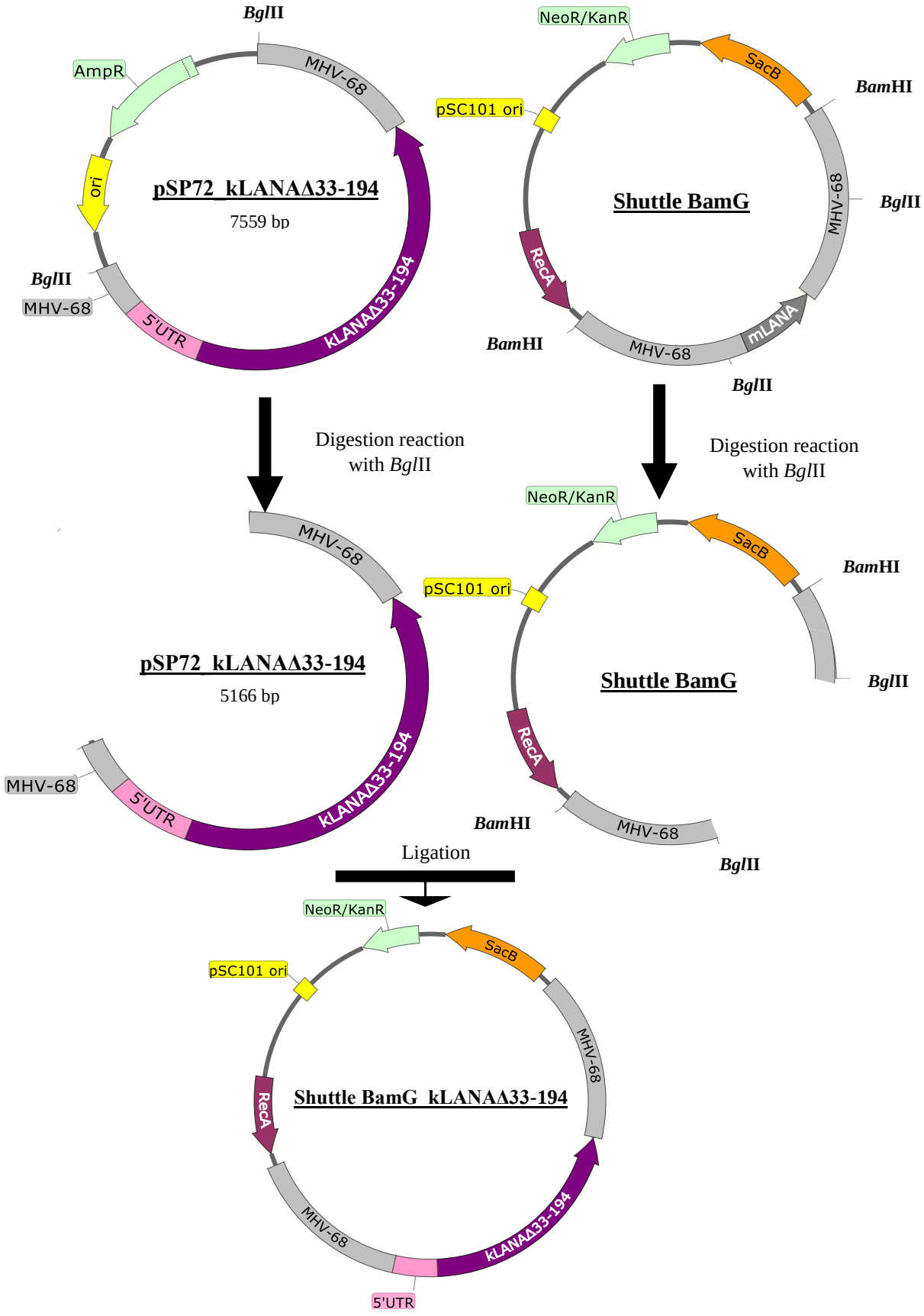


Figure S1. Schematic representation of cloning procedure described in Materials and Methods, section 3.2.1.2.

Table S1. Primers used for PCR, sequencing and real-time PCR.

Oligonucleotides	Sequence (5'-3')	Use
IMM_NS2	CATCACCCCAGGATCCCTCAGACGGGGATGATGATCC	PCR
IMM_NS3	GAGGATGGCGCCCCCGGGAATGCGCCTGAGGTCGGG	and
IMM_NS4	GTCTCCGGAAAGATCCACCCTGCGTAGCCTG	Sequencing
kLANAF_NruI	CGAATACCGCTATGTACTCAGAACATC	
IMM_AF2	AAACTGCAGATCCTTATTGTCATTGTCATC	
kLANAR_StuI	CCTGAGATAATTTTTTAAGTCCGTATG	
LANA_F1019	GTCCCTTACAGACAGATAGATGATTG	
IMM_TR11	CTGAGGGATCCTGGGGTGATG	Sequencing
IMM_T7	TAATACGACTCACTATAGGG	
IMM_TR1	AAAGAATTCAATCACCTTGGCATCC	
IMM_seq5	GAAACAAAACGTTGAGCATCC	
IMM_seq6	GTGCAAGATTATGGGCTCTTCCAC	
M9-F	GCCACGGTGGCCCTCTA	Real-time
M9-R	CAGGCCTCCCTCCCTTTG	PCR

Table S2. Primary antibodies used in western blot assay.

Antigen	Antibody	Specie	Dilution	Company
kLANA	LN53	Rat	1:500	Advanced Biotechnologies
mLANA	mAb 6A3	Mouse (ms)	1:250	PSimas lab (Pires de Miranda *et al*., 2013)
M3	M3	Rabbit (rb)	1:4000	PSimas lab (Jensen *et al*., 2003)
eGFP	α-eGFP	Mouse	1:1000	Sigma
Actin	Actin	Rabbit	1:1000	Sigma

Table S3. Secondary antibodies used in western blot assay.

Antibody	Specie	Dilution	Company
Anti-rat HRP	Goat	1:5000	Jackson Immunoresearch
Anti-ms HRP		1:5000	
Anti-rb HRP	Donkey	1:000	GE Healthcare

Table S4. Antibodies used in flow cytometry assay.

Antigen	Reactivity (fluorochrome)	Specie	Dilution	Company
CD19	APC-H7	Rat	1:400	BD Biosciences
GL7	eF660	Rat	1:200	eBiosciences
CD95	PE	Rat	1:800	BD Biosciences
CD16/32	Anti-CD16/32 (blocking solution)	Rat	1:100	BD Pharmingen